

# **The use of mRNA differential display in the analysis of somitogenesis in mouse embryos.**

**Rajeev Gupta**

A thesis submitted in partial fulfillment of the requirements of University College,  
London for the degree of Doctor of Philosophy

Medical Research Council  
Laboratory of Eukaryotic Molecular Genetics  
The National Institute for Medical Research  
The Ridgeway  
Mill Hill  
London NW7 1AA  
United Kingdom

**October, 1997**

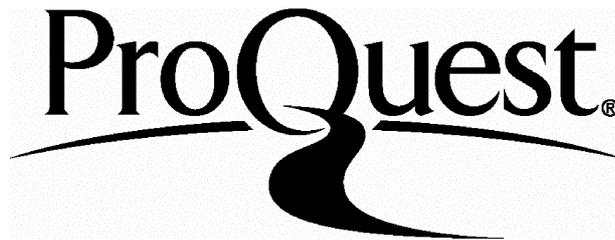
ProQuest Number: U643124

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 U643124

Published by ProQuest LLC(2015). 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

**For my parents, Indu and Subhash Gupta.**

## ABSTRACT

This thesis describes an investigation of differential gene expression, comparing dissected somites and the presomitic mesoderm of early mouse embryos. The method employed was mRNA differential display which uses arbitrarily primed PCR with short oligonucleotide primers to generate reverse transcription PCR fingerprints of the RNA species under study, with *in situ* hybridisation as a secondary screen of differential bands in the fingerprints. Initial studies revealed that although the original technique was able to generate reproducible fingerprints with nanogram quantities of material, there was a high rate of false positive clones, with multiple clones of identical size arising from any one band. Usually these clones, which are short, make poor *in situ* hybridisation probes.

In order to show the utility of the method, it was used to screen representative RNA made *in vitro* from two directionally cloned embryonic cDNA plasmid libraries. Using plasmid Southern blot hybridisation as a secondary screen, at least one clone generated from any single band was always truly differential.

The dual problems of high background and short products were overcome by using longer arbitrary primers at higher annealing temperatures in the PCR. Using this modified protocol, I have shown that the differential display principle can be used to compare the RNAs derived from singly dissected murine embryonic structures in the *in situ* hybridisation secondary screen.

I report the identification of a murine homologue of the *Drosophila* gene *tartan*, recently cloned from the adult brain, as a marker of early somites. Using a 5' RACE approach, I have isolated a novel relative of this gene which is expressed in a different somitic compartment.

# Table of contents

## CHAPTER 1

<b><u>Introduction</u></b>	<b>10</b>
1.1. Screening for differential gene expression in the mouse embryo	11
1.2. Conventional methods of screening for differential gene expression	12
1.2.1 cDNA Library construction from small amounts of RNA	12
1.2.1.1 PCR amplified cDNA libraries	13
1.2.2 Normalisation of cDNA libraries	14
1.2.2.1 Normalisation by PCR	14
1.2.2.2 Conventional methods of normalisation	15
1.2.3 Differential library screening and library subtraction	16
1.2.3.1 Differential screening	16
1.2.3.2 Library subtraction	16
1.2.3.3 Electronic subtraction	19
1.2.4 Random selection of clones from a mouse embryo library	19
1.3. Screening for novel genes using gene trap techniques	19
1.4. Differential screening using arbitrarily primed PCR	20
1.4.1 Differential display	21
1.4.2 Arbitrarily primed PCR	24
1.4.3 Theoretical considerations in differential display	24
1.4.3.1 Optimal PCR primer length	25
1.4.3.2 Displaying the entire mRNA population	26
1.4.3.3 The sensitivity of differential display and AP-PCR	26
1.4.4 Practical problems in differential display and potential solutions	28
1.4.4.1 Reducing false positive bands by modifying differential display	29
1.4.4.2 Alternative secondary screening strategies	30
1.4.5 The technical evolution of differential display	31
1.4.5.1 Template preparation	31
1.4.5.2 Primer strategies	31
1.4.5.3 Changes in gel technology	32
1.4.5.4 Fluorescent differential display using automated sequencers	33

1.4.6 Application of differential display to systems with limited RNA availability	33
1.4.6.1 Differential display at the level of a single cell - <i>in situ</i> transcription	35
1.5. Other novel methods of screening for differential gene expression	36
1.5.1 Domain specific differential display	36
1.5.1.1 Multiplex profiling	36
1.5.2 Fingerprinting of 3' end cDNA fragments	37
1.5.2.1 Differential 3' end cDNA profiling	37
1.5.2.2 Gene expression fingerprinting (GEF)	37
1.5.3 Combining subtraction and differential display	38
1.5.3.1 Representational difference analysis (RDA)	38
1.5.3.2 Suppression subtractive hybridisation (SSH)	41
1.5.4 Serial analysis of gene expression (SAGE)	41
1.5.5 Expression monitoring by hybridisation to high density cDNA arrays	46
1.6. Systematic comparisons of techniques for differential screening	47
1.7. Somite formation in the mouse embryo	47
1.7.1 The role of cell surface molecules in somitogenesis	51
1.7.2 The role of transcription factors in somitogenesis	52
1.8. Aim	53

-----

## CHAPTER 2

<b><u>Materials and methods</u></b>	<b>54</b>
2.1 Materials	54
2.1.1 Bacterial host cell strains	54
2.1.2 General chemical reagents and solutions	54
2.1.3 Synthetic oligonucleotides	62
Plasmid vector specific	62
Gene specific	63
Family specific	64
Differential display	64
Differential display band reamplification	65

Adapters	65
2.2 General techniques	65
2.2.1 Preparation and analysis of DNA and RNA	65
Extraction of nucleic Acids with phenol and chloroform	65
Precipitation of DNA with ethanol	66
Precipitation of DNA with isopropanol	66
Precipitation of DNA with polyethylene glycol (PEG)	66
Precipitation of in vitro transcribed RNA with ethanol	66
Quantitation of DNA and RNA in aqueous solution	66
Agarose gel electrophoresis of DNA	67
Agarose gel electrophoresis of RNA	67
Alkaline agarose gel electrophoresis of single stranded DNA	67
Extraction of DNA fragments from agarose gels	67
Southern blotting of plasmid DNA	68
Rapid alkaline Southern blotting of PCR products	68
Northern blotting of RNA from formaldehyde gels	68
2.2.2 <i>E.coli</i> , plasmids and bacteriophages	69
Preparation of competent bacterial host cells	69
Chemical transformation of competent bacterial host cells	69
Electro-transformation of commercially acquired competent cells	69
Preparation of bacteria for plating of $\lambda$ bacteriophage	69
Plating of $\lambda$ bacteriophage	70
Large scale plasmid DNA preparations from bacterial cultures	70
Small scale plasmid DNA preparations from bacterial cultures	71
2.2.3 General radiolabelling techniques	71
Preparation of radioactive size markers	71
Preparation of $\alpha^{32}\text{P}$ -labelled riboprobes for filter hybridisation	71
Preparation of $^{35}\text{S}$ -labelled riboprobes for <i>in situ</i> hybridisation	72
Random primer - labelling of DNA probes for filter hybridisation	72
Estimation of the specific activity of radioactively labelled probes	72
2.2.4 Filter hybridisation	73
Southern blots and colony lifts	74
Northern blots	74
2.3 Specific protocols	74
2.3.1 Collection of mouse embryos	74
2.3.2 RNA isolation and purification from mouse tissues	74
2.3.3 Purification of poly (A) <sup>+</sup> RNA from total RNA	75

2.3.4 In vitro synthesis of representative RNA from directionally cloned cDNA plasmid libraries	75
2.3.5 Synthesis of cDNA from dissected somites and presomitic mesoderm	75
2.3.6 Reverse transcription - PCR analysis of embryonic RNA	76
2.3.7 mRNA differential display (conventional protocol)	76
2.3.8 mRNA differential display (modified protocol)	77
2.3.9 Gel electrophoresis of differential display products	78
2.3.10 Elution and reamplification of DNA bands from native differential display gels	78
Conventional Protocol	78
Modified protocol	79
2.3.11 Cloning of reamplified differential display fragments	79
Bacterial transformation and PCR screening of recombinants	80
2.3.12 Library screening	81
Plasmid libraries	81
Bacteriophage libraries	81
Excision of Bluescript™ from Lambda ZAP® vectors	82
2.3.13 "Shotgun" cloning of full - length cDNA clones prior to sequencing	83
2.3.14 DNA sequencing	83
Radioactive sequencing of plasmid DNA	84
Automated sequencing	84
Template preparation for automated sequencing	84
Analysis of sequence data	84
2.3.15 Radioactive in situ hybridisation	85
2.3.16 Whole mount in situ hybridisation (Long protocol)	86
2.3.17 Whole mount in situ hybridisation (Rapid protocol)	88

-----

## CHAPTER 3

<b><u>The initial workup of mRNA differential display</u></b>	<b>90</b>
3.1 RNA preparation	90
3.2. Differential display	91
3.2.1 Reagents	91
3.2.2 Gel type	91

3.2.3 The effect of poly (A) <sup>+</sup> selection	91
3.2.3 The effect of RNA concentration	92
3.3. Cloning of differential display products	97
3.3.1 The use of the conventional reamplification protocol	97
3.3.2 A modified reamplification protocol	100
3.3.3 Cloning vectors	105
3.3.4 Screening of recombinant colonies	105
3.4. A comparison of RNA from 7.5 <i>dpc</i> and 8.5 <i>dpc</i> embryos	106
3.4.1 Sequence analysis	106
3.4.1.1 Differential bands	111
3.4.1.2 Non - differential bands	111
3.4.1.3 Up- or down- regulated species	112
3.4.1.4 Primer annealing sites	112
3.4.2 Database interrogation	112
3.4.3 Secondary screening	113
3.5 Summary	114

-----

## CHAPTER 4

### **Differential plasmid library screening using arbitrarily primed PCR** 116

4.1 Differential plasmid library screening by arbitrarily primed PCR	116
4.1.1 An outline of the library construction	117
4.1.2 Template preparation and differential display	117
4.1.3 Sequence analysis of differential bands	120
4.1.4 Secondary screening by plasmid Southern blot	120
4.2 Further analysis of differential clones	121
4.2.1 Isolation of full-length clones	121
4.2.2 In situ hybridisation analysis of full - length clones	124
4.2.3 “Shotgun” cloning for sequence analysis	125
4.2.4 Northern blot analysis	125
4.2.5 Clone e661	125

4.2.6 Clone m634	128
4.3 Summary	133

-----

## CHAPTER 5

### **The analysis of differential gene expression in mouse somites and presomitic mesoderm using mRNA differential display** **134**

5.1 A modified differential display protocol	134
5.2. A comparison of mouse somites and presomitic mesoderm	137
5.2.1 Tissue isolation and RNA preparation	138
5.2.2 RNA quality analysis and differential display	138
5.2.3 Secondary screening of differential clones	141
5.2.4 Clone S6	141
5.2.5 Clone P4	142
5.3. Refinements in template preparation	145
5.3.1 Differential display of individually dissected embryonic structures	145
5.3.2 Clone S53	148
5.4 Isolation of homologues of NLRR-1	153
5.4.1 A 5'RACE strategy	153
5.4.2 Analysis of LRR-4	156
5.5 Summary	161

-----

## CHAPTER 6

### **Discussion** **162**

6.1 The generation of reverse transcription PCR fingerprints	163
6.1.1 Template preparation	163
6.1.2 The reproducibility of display patterns	164

6.1.3	The complexity of the display patterns	165
6.2	Analysis of differential bands	165
6.2.1	False clones	166
6.2.2	Identification of true positive clones	167
6.2.3	Differential library screening by arbitrarily primed PCR	167
6.2.4	Secondary screening of differential clones	168
6.3	A modified differential display protocol	168
6.3.1	The elimination of false clones	168
6.3.2	The generation of long clones	172
6.3.3	The frequency of true positive bands	172
6.3.4	Primer annealing sites	172
6.4	The abundance of clones isolated by arbitrarily primed PCR	172
6.5	A family of transmembrane molecules expressed in the somite	173
6.5.1	The leucine rich repeat motif	174
6.5.1.1	Evolutionary relationships of LRR genes	174
6.5.1.2	The crystal structures of LRRs	175
6.5.1.3	The functions of LRR containing proteins	175
6.5.2	Neuronal leucine rich repeat proteins	176
6.5.2.1	The isolation of novel NLRR homologues from the mouse embryo	177
6.5.2.2	Homologues of the NLRR genes in drosophila	177
6.6	Conclusions and future directions	178
6.6.1	Further applications of differential display	178
6.6.2	Further studies of the somitic LRR genes	179

-----

**Acknowledgements**

**180**

-----

**Bibliography**

**181 - 190**

## Index of figures

Figure 1.1	22
Figure 1.2	39
Figure 1.3	42
Figure 1.4	44
Figure 1.5	49
Figure 2.1	93
Figure 2.2	95
Figure 2.3	98
Figure 2.4	101
Figure 2.5	103
Figure 2.6	107
Figure 2.7	109
Figure 3.1	118
Figure 3.2	122
Figure 3.3	126
Figure 3.4	129
Figure 3.5	131
Figure 4.1	135
Figure 4.2	139
Figure 4.3	143
Figure 4.4	146
Figure 4.5	149
Figure 4.6	151
Figure 4.7	154
Figure 4.8	157
Figure 4.9	159
Figure 5.1	170

# CHAPTER 1

## Introduction

A characteristic feature of the embryology of all vertebrates is the formation of serially repeated structures in the trunk. This process of segmentation, manifested by the appearance of somites occurs soon after germ layer formation and is remarkably conserved both numerically and at the morphological level.

Segmentation in the trunk impacts directly upon the morphogenesis of the body axis and the somites are known to impart a periodic restriction on the position of various adult anatomical structures *e.g.* the roots of the spinal cord (Trainor and Tam, 1994). Superimposed upon this globally segmented pattern, each somite has two polarities of cell differentiation. The rostro-caudal axis, established prior to somite formation and the dorsoventral pattern that is established with somite maturation.

The last few years have witnessed an enormous increase in understanding of the many genes expressed in the subpopulations of somites and also of those genes thought to specify positional identity along the rostro-caudal axis of the embryo. It is likely that regional identity is conferred by combinatorial expression of *Hox* genes and that this in turn plays some role in determination. Much less is known about about the earlier steps of somite formation and how the subcompartments of the mature somite actually arise.

A systematic search for the genes and gene products involved in these processes is an essential part of progress in their understanding. This will allow greater insights into the very earliest steps of cellular determination and differentiation, which are critical questions in biology and medicine.

Strategies that compare gene expression in different tissues of the developing embryo are a useful but technically challenging approach for addressing such questions. Biologically interesting comparisons may involve adjacent somites on either side of a *Hox* gene boundary of expression, different halves of an individual somite or the rostral and caudal ends of the presomitic mesoderm. mRNA differential display is a novel technique for performing such comparisons that may overcome the difficulties in the conventional methods available

## **1.1. Screening for differential gene expression in the mouse embryo**

The search for the gene products that are involved in patterning, cellular determination and migration in the developing mammalian embryo is an area of extensive investigation. Although many critical proteins in embryogenesis are ubiquitously expressed, at least some of the genes that are uniquely expressed in the different stages of development have yet to be isolated. Many of the known examples of such genes in higher animals have been found either by homology to species isolated from lower animals or by the study of well characterised developmental mutations.

There are relatively few methods whereby one could find such gene products without any *a priori* information regarding their structure, based purely on their pattern and/or timing of expression. The very fact that most of the known genes that are expressed in spatially and temporally regulated patterns in the mouse embryo have been cloned by homology speaks for the technical difficulties inherent in, and the relative inefficiencies of, these procedures. In principle, a large scale mutagenesis screen in mice would provide the most complete information regarding the genes involved in the various processes of development but the costs of such an enterprise are very high.

The murine embryologist faces one major problem when he or she attempts to study differential gene expression. There is a relative paucity of tissue available from which RNA can be made for the purposes of northern blotting, library construction,*etc.* and so innovative technical approaches have been developed by necessity with the emerging science of molecular embryology in order to circumvent this drawback. Classical methods for differential screening are based on library construction. They are labour intensive and often require considerable amounts of starting material. Differential display and related methodologies rely upon PCR-based amplification of entire mRNA populations followed by a direct comparison of the prevalences of the resulting RNAs in the tissues being studied. Since its inception in 1992, differential display has offered a potentially tantalising solution to the fundamental problem of limited tissue availability in differential screening. As a primary screen it can be performed with nanogram quantities of RNA and is outwardly not labour intensive, requiring only a single step. However, despite some 300 scientific publications at the time of writing, there have been no reports of its successful application in studies of mouse embryos. This is an indication of the technical problems in the method which

are paradoxically enhanced in this area of research. In this introduction, I shall outline the traditional library based methods of screening for differential gene expression as applied to systems where there is limited access to RNA, gene trap strategies and then focus on differential display and related PCR-based techniques for gene discovery. Theoretical and technical considerations in the evolution of differential display methodologies will also be considered. Finally I shall briefly review what is known of the molecular biology of somitogenesis in the mouse and outline the aims of the experiments presented subsequently.

## **1.2. Conventional methods of screening for differential gene expression**

These techniques have been based upon the prerequisite for cDNA library construction from the tissue(s) of interest allied with either differential screening or subtraction in order to isolate clones specific to one tissue or the other. Their application in systems where RNA is limiting has been reported though with varied success

### **1.2.1 cDNA Library construction from small amounts of RNA**

A particularly rare mRNA species may comprise only one molecule in  $10^6$  of a preparation of poly (A)<sup>+</sup> RNA (Davidson and Britten, 1979) and the main problem in creating any library, no matter how much RNA one has to start with, is the production of the huge population of clones necessary to ensure that it contains at least one copy of each clone of interest (Sambrook *et al.*, 1989). There are now many different strategies for conventional cDNA library construction. Each new method or modification is aimed to overcome some problem experienced with older procedures. Typical protocols require a minimum of 1 $\mu$ g poly (A)<sup>+</sup> RNA although the high cloning efficiency offered by electrotransformation has permitted the construction of plasmid based cDNA libraries from as little as 10ng starting poly (A)<sup>+</sup> RNA (Rothstein *et al.*, 1993). A single mouse embryo at the stage of implantation may contain as little as 50pg poly (A)<sup>+</sup> mRNA (Clegg and Pikó, 1983). Nevertheless, libraries from unfertilised eggs, two cell, eight cell and blastocyst stage mouse embryos have been made conventionally (for example, Rothstein *et al.*, 1992). Rothstein and coworkers used between 600 and 13500 whole embryos from the various stages under study in order to generate between 45ng and 175ng poly (A)<sup>+</sup> RNA for the production of representative plasmid libraries with approximately  $10^6$  clones in each. The generation of such libraries from whole embryos provides a relatively crude resource with which to study later events in embryogenesis where questions regarding positional identity or cell type are pertinent.

Only one attempt to generate conventional libraries from dissected individual tissues of mouse embryos has been published (Harrison *et al.*, 1995). These authors made dissections of the three germ layers and primitive streaks from 265 7.5 *dpc* mouse embryos and used total RNA to produce (again via electrotransformation) plasmid based libraries with approximately  $10^6$  clones each. This is a considerable feat of fine dissection and has generated libraries that were subsequently shown to contain novel tissue specific genes which have the expected patterns of expression on *in situ* hybridisation (S.Dunwoodie, Laboratory of Mammalian Development, NIMR). It is an undertaking that would not be taken lightly in any developmental laboratory. The production of libraries is often prone to mishaps and very often more than one attempt is required. If small amounts of RNA are available then painstaking dissections become the rate limiting step. Furthermore, many workers would be interested in finding genes expressed differentially between even smaller structures than dissected germ layers, *e.g.* adjacent somites or rhombomeres.

#### **1.2.1.1 PCR amplified cDNA libraries**

There have been several reports of methods for cDNA library synthesis from a few or single cells (Gurr *et al.*, 1991; Brady and Iscove, 1993; Lambert and Williamson, 1993; Luqmani and Lymboura, 1993; Dulac and Axel, 1995). These have all been variations on the same idea of PCR amplification of the small amounts of cDNA made during first - strand synthesis. The basic method has gradually evolved into a single tube reaction in which all steps from RNA extraction to PCR can be performed without the need for multiple transfers or precipitation steps (Brady and Iscove, 1993). Oligo dT-primed cDNA synthesis is initiated *in situ* on the total RNA of between 1 and 40 cells that have been lysed in mild detergent (0.5% (v/v) NP40). Usually both avian myeloblastosis virus and Moloney murine leukaemia virus reverse transcriptases are used in order to avoid any sequence specificity attributable to either one alone. After this the cDNA is homopolymer tailed with dATP and the resulting mixture amplified by PCR using a single primer (up to a 60mer oligonucleotide) that has a long 3' dT stretch. The resulting amplified cDNA can be cloned into phage or plasmid vectors. Using this approach, representative libraries containing up to  $10^5$  clones can be generated from single cells as assessed by screening with probes for known markers. Libraries with similar numbers of individual clones have been made from adjacent normal and abnormal breast lobules in a haematoxylin/eosin stained  $6\mu\text{M}$  frozen surgical specimen (Luqmani and Lymboura, 1993). However, these libraries generally have small mean insert sizes. The PCR step will tend to give progressively shorter mode fragment sizes with each

cycle and the library will thus become biased against longer full length cDNA species. In order to bypass this problem and retain the utility of the library for the purposes of screening, the initial reverse transcription step can be prematurely terminated in order to purposely generate "3' UTR" cDNA libraries. Differential clones between such libraries constructed from individual cells in a tissue can be used to screen more complete libraries made from whole tissues. The utility of these methods in embryological systems has yet to be seen. Recently however Bevan and coworkers (Bevan *et al.*, 1996) have constructed PCR amplified libraries from the neural crests of quail embryos from different stages and also from neural crest derivatives in older embryos after their differentiation into melanocytes, adrenal medullary cells and Schwann cells. Dissection in an avian embryo is considerably easier than in a mouse embryo and it is not clear as yet how many novel spatially restricted genes will be cloned from these libraries.

### **1.2.2 Normalisation of cDNA libraries**

Reassociation - kinetic analyses have shown that the mRNAs of a typical cell are distributed into 3 main frequency classes the abundances of which should be maintained in a representative cDNA library ( Bishop *et al.*, 1974). Rare species may be present at only 1 or 2 copies per cell whereas others will have up to 5000 copies (Bishop *et al.*, 1974; Davidson and Britten, 1979). The very rarest cDNAs in a representative library may have a frequency of  $10^{-6}$  or even lower depending on the sequence complexity of transcripts in the source tissue. Workers who have developed technology for gridding individual clones in a library and who have been interested in high throughput single pass sequencing of random clones from libraries (Adams *et al.*, 1992; Soares *et al.*, 1994) have found it advantageous to use libraries that have been normalised (*vide infra*). Such libraries have reduced representation of the (redundant) clones of abundant mRNA species with relative enrichment for clones of rare transcripts.

#### **1.2.2.1 Normalisation by PCR**

There are several ways by which normalisation can be achieved. The PCR-based methods use the inherent normalising properties of PCR, although these properties of PCR were not initially acknowledged by the authors. During a PCR where substrates are not limiting, the rate of synthesis of product will rise exponentially until a plateau is reached at which point its amount is such that at the annealing step, the single strands of the template are more likely to re-anneal than bind with fresh primer. Any PCR performed on a heterogeneous template, be it with random primers as in

differential display or linker primers as in PCR directed library construction will lead to the relative enrichment of low abundance species as more cycles are completed (Mathieu-Daudé *et al.*, 1996b). This is because the plateaux for the synthesis of common products are reached before those for rarer species; thus at each successive cycle, the primers in the reaction are more likely to anneal to rarer templates and so differences *within* the starting mixture of templates are reduced. This has implications for both differential display and cDNA library construction, in that relative differences in transcript abundance *between* samples can be hidden and contaminants can be enriched. However, the degree to which a template in any such PCR is actually normalised is unknown. Ko (1990) described the construction of an “equalised” library in which full length double stranded cDNA was sonicated to a mean size of 200 - 400 bp, the fragments polished and a single stranded linker ligated to each end. A primer complementary to the linker was then used to PCR amplify the cDNA after which it was denatured and allowed to reassociate. Single stranded cDNAs were purified by hydroxylapatite (HAP) column chromatography, re-amplified and cloned. Several rounds of this could be performed before cloning of the cDNA and colony hybridisation of the final libraries showed a 500-fold reduction in the abundance variation of clones representing various genes.

### 1.2.2.2 Conventional methods of normalisation

These are based upon hybridisation and antedate the PCR based approaches. Current approaches have sought to generate full length normalised libraries (Soares *et al.*, 1994; Bonaldo *et al.*, 1996). They have been based on directionally cloned plasmid libraries. The basic principle has been to convert the library into a single stranded circular form either *in vitro* by the use of phage FI endonuclease and exonuclease III or *in vivo* by M13 helper bacteriophage superinfection. Then the single stranded DNA is hybridised with single stranded “blocking” DNA or RNA oligonucleotides derived from the libraries themselves and purified single stranded circles are either used for a second round of normalisation or converted into partial duplexes (usually by random priming) and retransformed as a normalised library. The blocking oligonucleotides can be made by a variety of approaches; because HAP chromatography is so specific in separating single stranded DNA from partial DNA duplexes or RNA/DNA hybrids they do not have to be full length representations of library clones. The commonest way of generating them is by *in vitro* transcription from the endogenous plasmid promoters, although heat denatured whole library PCR amplicons (made by using vector-specific primers) and single stranded DNA from limited primer extension reactions performed on single stranded library have also been used (Soares *et al.*, 1994; Bonaldo *et al.*, 1996). Multiple rounds of

normalisation can be performed; the secret to success in this endeavour lies not only in the hybridisation step but also in the HAP chromatography, where some hybrids may melt in the column buffer. The conditions for these steps have been carefully worked out (Soares *et al.*, 1994).

### **1.2.3 Differential library screening and library subtraction**

#### **1.2.3.1 Differential screening**

Once a library or libraries have been constructed, the simplest approach to isolating clones specific to one particular tissue or tissue state and not another is to perform a simple differential hybridisation screen (*e.g.* Nedivi *et al.*, 1993; Sasai *et al.*, 1994; Dulac and Axel, 1995). In this strategy, library filters are hybridised with radiolabelled probes generated from either RNA of the tissue(s) under study or other libraries. Clones which hybridise with only with the probe of interest are picked for a secondary screen. Whilst in principle this is attractive there are problems. Firstly, probes made in this way from heterogeneous sources (even from normalised libraries) must be labelled to very high specific activities and used at very high concentrations and even then the amount of label incorporated into rare species is low. In order to enrich such complex probes for relatively low abundance species, it is possible to reduce the amounts of abundant and non-differential RNA by subtractive hybridisation (Vitek *et al.*, 1981; Mohn *et al.*, 1990). Here RNA from tissues not of interest is bound to an oligo-dT cellulose column and used to hybridise to the cDNA of interest. Non-hybridised cDNA is thus enriched for species specific to the tissues of interest. In a more recent variation on this theme, biotinylated first strand cDNA is generated from one tissue and immobilised onto magnetic streptavidin beads and hybridised with poly (A)<sup>+</sup> enriched RNA from a second tissue (of interest). After magnetic separation of the RNA/cDNA hybrids, the RNA left in solution can be used as an enriched probe (Aasheim *et al.*, 1994). Protocols for generating radiolabelled cDNA from RNA directly require at least 100ng mRNA (Sambrook *et al.*, 1989) and subtractive enrichment uses ten times this amount. This is an obvious limitation for the mouse embryologist.

#### **1.2.3.2 Library subtraction**

In order to increase the yield of differential screening experiments many libraries are now subtracted. This is a procedure similar to library normalisation except that instead of being normalised against itself, the library is normalised with respect to a

library or RNA made from a different tissue (Bonaldo *et al.*, 1996). Consequently, subtracted libraries should be enriched for clones unique to or expressed at much higher levels in the tissues from which they were made. The principle is to generate single stranded DNA (target) from the library to be subtracted and hybridise it with DNA or RNA fragments (driver) made from a different library. Usually tens of micrograms of driver material are required and so RNA derived from tissues is rarely used. Classically, non-hybridising target DNA is purified after HAP chromatography and either re-subtracted or used directly as an enriched library (Rubenstein *et al.*, 1990). In general many workers have found HAP chromatography to be difficult and unreliable because of problems with hybrids melting in HAP column buffer. Consequently, methods that employ biotin and streptavidin have come to the fore (Swaroop *et al.*, 1991). The streptavidin can be complexed to a solid phase (Swaroop *et al.*, 1991) or more often to paramagnetic particles (Luqmani and Lymboura, 1993) in order to facilitate the easy separation of biotinylated species from enriched driver DNA or RNA. The driver DNA can be biotinylated and unhybridised target DNA separated from hybrid molecules and unhybridised biotinylated DNA using streptavidin. Subtraction of the early mouse embryo libraries of Rothstein *et al.*, (1992) was achieved by *in vitro* transcription of biotinylated excess RNA from the driver libraries and unlabelled antisense RNA from the target libraries. After hybridisation, all biotinylated species (hybrids and excess driver RNA) were removed with streptavidin and the remaining (unbiotinylated) enriched target RNA was used for reverse transcription and cDNA library construction. The original libraries were preserved and could be used for the isolation of full length clones if required. Library subtraction with RNA or DNA derived from another library runs the risk of polylinker homology between target DNA and driver DNA or RNA leading to inappropriate subtraction. Using a modification of an earlier protocol (Rubenstein *et al.*, 1990), Harrison and coworkers circumvented this problem in their production of subtractive libraries from embryonic germ layers (Harrison *et al.*, 1995) by making a second driver library from purified total cDNA excised from each original plasmid library. From this second library, biotinylated driver RNA was made and used to subtract single stranded target library DNA. Directional tag PCR subtraction (Usui *et al.*, 1994) involves the subtraction of *in vitro* transcribed RNAs from target and driver libraries, HAP column separation of unhybridised target and driver RNA, followed by specific reverse transcription - PCR amplification of target RNA using primers specific to sequences in the target library vector polylinker.

There are now several newer protocols for the construction of subtracted libraries which have been aimed at reducing the technical difficulties encountered in conventional protocols. That none of them has been widely adopted is an indicator of

the general difficulties with each of these methods. They attempt to enrich the target RNA prior to cloning but in some cases require considerable amounts of RNA. Chemical cross-linking subtraction (Hampson *et al.*, 1992) utilises 2,5 diaziridinyl - 1,4 - benzoquinone to chemically cross-link driver RNA and a cDNA target. The remaining single stranded cDNA can then be used without further purification for probe production by random priming. However this method requires several micrograms of target cDNA. Hara *et al.*, (1991) poly (A)<sup>+</sup> enriched and reverse transcribed driver RNA using oligo(dT)<sub>30</sub> latex beads and used the resultant cDNA to enrich target RNA. After several rounds of subtraction nanogram quantities of the remaining driver were used to make a PCR amplified cDNA library. This method also required microgram quantities of RNA prior to enrichment.

Ghosh (1996) has reported a PCR-based method for the generation of subtracted libraries from as little as 200ng poly (A)<sup>+</sup> enriched RNA. The system under study was the protozoan *E.crassus* as it left the vegetative state (where tissue availability is not a problem) and entered macronuclear development (where RNA extraction is difficult). Here approximately 200ng target RNA from developing cells was reverse transcribed with an adapter primer and a second anchor adapter primer was ligated to the 3' end of the cDNA using RNA ligase. After hybridisation with a large excess of biotinylated driver RNA made from a directionally cloned vegetative cell plasmid library, purified target cDNA was PCR amplified for cloning as a subtraction plasmid library. A potentially interesting method for the embryologist is enzymatic degradation subtraction (Zeng *et al.*, 1994). Double stranded target cDNA, which could be generated by PCR from limited amounts of tissue is digested by Klenow enzyme and the recessed termini produced filled in with phosphorothioate nucleotides. The modified target DNA is heat denatured and hybridised with an excess of denatured PCR-amplified driver cDNA. The resulting mixture contains target-target homohybrid species, driver-tester heterohybrids and driver-driver homohybrids which are digested with exonuclease III, an enzyme inactive on phosphorothioate containing DNA. The only double stranded species left after this procedure are double stranded enriched target molecules which are then amplified by a few cycles of PCR prior to cloning. Finally, Hakvoort and coworkers (Hakvoort *et al.*, 1996; T. Hakvoort, 1996, personal communication) have developed a rapid method for the production of an enriched cDNA library by the utilisation of triple helix affinity capture, which relies on the RecA protein-mediated stabilisation of triple helical structures formed between single stranded and double stranded DNA molecules on account of their sequence similarity. They generated triple helices between a cDNA plasmid library to be enriched and digoxigenin labelled single

stranded driver cDNA. After removal of the complexes using antidigoxigenin paramagnetic beads, they were left with an enriched plasmid library.

### **1.2.3.3 Electronic subtraction**

This is an experimental concept employed in the computer analysis of random clone sequences from normalised differential cDNA libraries that aims to predict which clones will be differential by some biochemical screen, *e.g.* Northern analysis. If first pass sequencing is performed on for example, 10000 random clones from two libraries, computer algorithms are being developed to select which clones are represented in one library over the other at whatever ratio or stringency is required. At the moment, these are hampered by the need to sequence virtually the whole library in order to generate statistically meaningful data. (Mark Erlander, RW Johnson Pharmaceutical Research Institute, San Diego, Personal communication).

### **1.2.4 Random selection of clones from a mouse embryo library**

Given the inherent difficulties in differential screening and subtraction of libraries, some workers have suggested that having a quality library derived from mouse embryo RNA may be all that is necessary to isolate novel genes with interesting differential patterns of expression (Bettenhausen and Gossler, 1995). These workers made a 10.5 *dpc* mouse whole embryo cDNA library and randomly selected 71 cDNA clones for sequencing. *In situ* hybridisation analysis was performed on 41 of these encoding novel mouse genes and 10 gave spatially restricted patterns of expression. The remainder were ubiquitously expressed or gave no *in situ* hybridisation signal. One of these was a murine gene related to *Drosophila delta* (Bettenhausen *et al.*, 1995). This outwardly sensationally high “hit rate” of 24% of randomly selected novel clones giving interesting patterns of expression has yet to be matched in the literature. Such strategies applied to libraries made from dissected embryonic germ layers have yielded a less than 1% hit rate (S.Dunwoodie, NIMR, personal communication).

### **1.3. Screening for novel genes using gene trap techniques.**

Although not strictly a strategy for *differential* screening, gene trap approaches have been used in mice to find genes that are important in mammalian development and the genes that they drive (Joyner, 1991). They represent the nearest method to the large scale mutagenesis screens that have been carried out in lower organisms for

mutations affecting development (Hill and Wurst, 1993) and they allow for the simultaneous identification and mutagenesis of novel genes.

The method requires the construction of a gene trap vector which contains a single splice acceptor site upstream of an assayable reporter gene (typically the *E. coli* LacZ gene). This is allowed to randomly insert into the genome and when inserted into an intron, a fusion gene of the gene at the target locus and lacZ is transcribed which can be easily monitored in the developing embryo. The vectors also contain a selectable marker for antibiotic resistance.

The gene trap vector is either transfected (Joyner, 1991) or infected retrovirally (Brenner *et al.*, 1989) into ES cells which are used to generate chimaeric blastocysts and random reporter gene insertions are assayed in the resulting embryos.

As a method for identifying novel genes in embryogenesis, it is basically undirected though various modifications have been applied in order to focus on particular gene types. Skarnes *et al.*, (1996) have attempted to capture genes encoding secreted molecules by inserting a transmembrane domain upstream of the LacZ gene but just after the splice acceptor site in the gene trap vector. Only insertion of the vector into the introns of genes carrying signal sequences will result in fusion proteins where the reporter is oriented towards the cytosol and therefore active. Another approach has been to screen transfected ES cells with retinoic acid prior to chimaera generation in order to find retinoic acid responsive genes.

These technologies require considerable resources and expertise in order to be carried out properly. However they do rely upon techniques that are well developed in many mammalian development laboratories and do not require the extraction of RNA.

#### **1.4. Differential screening using arbitrarily primed PCR**

When PCR is performed on a complex template such as the total cDNA of a cell or genomic DNA using a primer or primers of arbitrarily chosen sequence, then at sufficiently low temperatures annealing occurs and DNA polymerisation is initiated. Although the initial priming event is sequence dependent the match of primer and template is usually only seen at the 3' terminal 6-8 bases of the primer (McClelland *et al.*, 1995). However after the first two cycles of PCR the products of this initial DNA polymerisation will have primer sequences at both ends and can be amplified by PCR. The most prominent products visualised at the end of the PCR will derive from templates which have the most efficient pairs of priming sites separated by the most easily amplified sequence. Each heterogeneous template will lead to its own specific

pattern of amplicons based upon the sequences of which it is composed. Some amplicons will be peculiar to that template and represent sequences or genes specific to it. This is the principle of a family of differential screening protocols based on PCR, of which differential display is the most widely used.

#### **1.4.1 Differential display**

The term differential display was coined by Liang and Pardee (1992) to describe a novel method for differential cDNA screening and cloning. These workers are tumour biologists who had been studying differences in gene expression in breast carcinoma lines using two dimensional protein electrophoresis. Very often it was possible to identify a species seen on the gel as being differential but almost impossible to purify enough of it to acquire meaningful sequence data. They had tried to develop a two dimensional RNA fingerprinting technique to circumvent the problem but this was unsuccessful as a result of the huge amounts of RNA required. Consequently they changed direction and turned to random PCR directed cDNA amplification as a way of generating an array or fingerprint of cDNAs specific to a given cell type. A comparison of these arrays from two different RNA sources would lead to the identification of cDNA species specific to one and not the other.

The basic method of Liang and Pardee is outlined in Figure 1.1. mRNA from a given tissue is reverse transcribed with a short oligo (dT<sub>11</sub>) primer with an anchor of two bases (*e.g.* AG) at its 3' end. Priming preferentially occurs at the 5' ends of the poly (A) tails of the mRNA species and selectively in sequences that end in 5' UC-poly (A). (By probability, this primer will recognise one in twelve of all mRNAs as there are 12 different combinations of the last two bases, omitting T as the penultimate base). PCR is then performed on this subset of the total first strand cDNA using the reverse transcription primer and an arbitrarily chosen 10-mer oligonucleotide primer. If the PCR annealing temperatures are low enough (38°C - 42°C is the generally accepted range) then the 10-mer primer should anneal at random points from the start of the poly (A) tail depending on the sequences in the cDNA population.

Consequently, the PCR will generate a series of products that, if radiolabelled, can be viewed on a polyacrylamide gel as a cDNA fingerprint or display that is unique to the starting RNA, the primers used and the PCR conditions employed. If PCR products from identical reactions with two different starting RNAs are compared in adjacent lanes of the gel, then it can be envisaged that species or bands in the fingerprint that are shared might represent common transcripts and those unique to one or the other are representative of specific genes. These can be easily eluted from the gel,

### **Figure 1.1**

Schematic representation of the generation of a reverse transcription PCR fingerprint using differential display. First strand synthesis is performed on one twelfth of the mRNA species using a short dT primer that has a two base 3' anchor (in this case GA). A low stringency PCR is then performed on the cDNA using the reverse transcription primer and a single arbitrary 10mer. This should anneal at random positions along the cDNA during second strand synthesis, thereby giving an array of products which when size fractionated by gel electrophoresis will provide a PCR fingerprint unique to the starting cDNA.

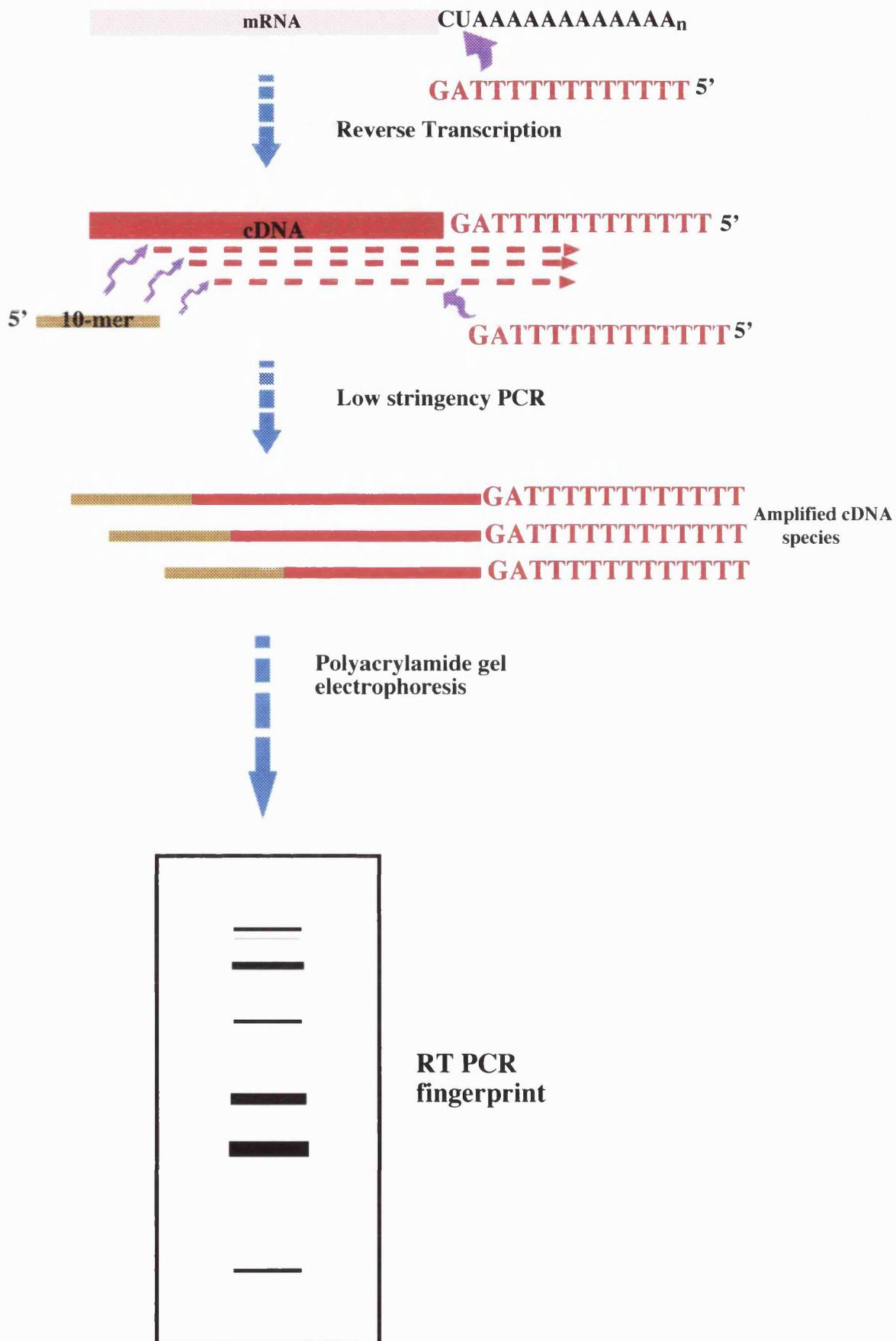


Figure 1.1

reamplified, cloned and used in some secondary screening process. It is possible that if a full range of reverse transcription primers is used with a large enough panel of random primers a reasonable percentage of the estimated 15000 genes expressed in a given cell could be displayed using this approach (Guimarães *et al.*, 1995b).

#### **1.4.2 Arbitrarily primed PCR**

At the same time as Liang and Pardee were developing differential display, other workers interested in the rapid identification and classification of bacteria had developed methods for reproducibly PCR fingerprinting genomic DNAs from various *Staphylococcus* strains (Welsh and McClelland, 1990). Initially, they had selected single arbitrarily chosen 24-mers that had been used for other purposes and their PCR parameters involved two cycles of low stringency annealing at 40°C followed by 40 cycles of high stringency annealing at 58°C - 65°C. Using the same polyacrylamide gels as Liang and Pardee, they observed reproducible species and strain specific bands from the cells of single bacterial colonies that had been picked and suspended directly into PCR buffer. This method was termed “Arbitrarily Primed PCR” (AP-PCR) and it was soon extended to various inbred mouse strains where they found that with only 3 primers used singly and in combination, they could generate on average 50 bands per PCR that were reproducible and strain specific in pattern (Welsh and McClelland, 1991).

Welsh and McClelland were able to extend their technique to RNA derived from the same adult tissues in different mouse strains (Welsh *et al.*, 1992) and independently developed analogous technology to differential display, where purified differential AP-PCR bands could be cloned and shown to be tissue and/or strain specific .

#### **1.4.3 Theoretical considerations in differential display**

When first published, the method predicted the generation of small PCR amplicons (up to 500 bases) that would be directional and biased to the 3' ends of genes. These amplicons represented single gene fragments and once cloned could be used in a secondary screen. It was quickly shown that differential display gave a reproducible fingerprint of cDNAs for any given RNA with as much as 2µg or as little as 2ng RNA (Liang *et al.*, 1993). As will be discussed later, it was rapidly observed that the method as originally proposed gave a high false positive rate as judged by the occurrence of bands that did not give rise to differentially expressed clones. In addition to this major problem, the clones that were obtained were generally small

and it was often hard to use these isolate full length species corresponding to differential genes (McClelland *et al.*, 1995) in conventional library screens.

Certain other of the initial predictions regarding the clones obtained have been found to be incorrect by practitioners of the technique. A significant number of amplicons arises by priming with only one of the two oligonucleotides (Guimarães *et al.*, 1995 b). Some primers are “weak” for PCR using some cDNA templates but not others. The clones isolated can reside anywhere along the full length of the genes from which they derive and they may contain only coding sequence. Neither of these findings should detract from the general principle of the method and its potential utility.

#### **1.4.3.1 Optimal PCR primer length**

Liang and Pardee found that arbitrary primers of 10-mer size gave on average 50 - 100 bands per display track and that low amounts of dNTPs in the reverse transcription step (20 $\mu$ M) and the PCR (2 $\mu$ M) were necessary for labelling the amplicons to high enough specific activity for their autoradiographic visualisation. Such low concentrations also improved the specificity of the PCR amplification. Although in principle, shorter arbitrary primers would prime more frequently, in practice this was not the case as no products were seen with them (Liang *et al.*, 1995). Most workers who have used differential display report that mismatches occur between the sequences of cloned genes and the arbitrary primer sequence (Liang *et al.*, 1993). They tend to occur at the 5' end of the primer and it is thought that the four most 3' bases of a 10-mer primer determine its annealing specificity (Liang *et al.*, 1992). The effect of using one or both primers in a differential display reaction has been investigated and it has been shown that the reverse transcription primer is only effective in altering the pattern of bands in a gel if used in at least a four-fold molar excess over the arbitrary primer for most arbitrary primers (Miyashiro *et al.*, 1994). This indicates its relative weakness as a PCR primer under the standard reaction conditions. The experience with AP-PCR has shown that longer primers will give rise to similar numbers of bands per reaction (Welsh *et al.*, 1992). Some workers have successfully used longer primers with 5' extensions to perform reverse transcription and differential display PCR using cycling parameters similar to those for AP-PCR (Ayala *et al.*, 1995; Diatchenko *et al.*, 1996a). The bands seen are reported to be easier to reamplify and clone and it is claimed that the PCR products are on average longer (Sokolov and Prockop, 1994). The 5' extensions used in these primers can have useful functions such as bacterial RNA polymerase promoters (Eberwine *et al.*,

1992), potentially useful for the rapid production of riboprobes from reamplified fragments, or restriction sites useful in cloning (Liang *et al.*; 1994).

#### **1.4.3.2 Displaying the entire mRNA population**

Theoretically, all expressed genes in a cell should be detectable using differential display. The representation of each message in a gel depends upon the efficiency of primer binding to it, the nature of the sequence between the primers, its amenability to amplification with *Taq* polymerase and also the abundance of the message. The number of primer combinations required to display all of the estimated 15000 transcripts in a given cell has been variously calculated at between 240 and 312 depending upon the number of bands per lane that individual workers see, which corresponds to approximately 75 arbitrary primers in conjunction with the 12 anchored reverse transcription primers (Guimarães *et al.*, 1995b; Liang *et al.*, 1995). In the hands of some workers this can correspond to as few as 26 primers (Bauer *et al.*, 1993). Bauer *et al.*, (1993) also made a formal mathematical prediction of how many bands should be generated by a single primer and unsurprisingly predicted that shorter primers would give more bands. However as Liang and Pardee had shown earlier, priming as assayed by the presence of visible PCR products hardly occurs with oligonucleotides less than eight bases long.

In practice, the number of primer combinations required to generate 15000 bands is dependent upon the PCR conditions employed, including the thermal cycler and the PCR reaction tubes used. By increasing the concentration of the arbitrary 10-mer from 2 $\mu$ M to 30 $\mu$ M and reducing the annealing temperature from 40°C to 32°C, a 25% increase in the total number of bands *per* lane has been seen (Guimarães *et al.*, 1995b). Increasing the concentration of magnesium in the PCR buffer from the 2.5 $\mu$ M used by Liang and Pardee to 4 $\mu$ M will also increase the likelihood of degenerate priming in the first rounds of the PCR (Welsh *et al.*, 1992). Most workers who have subsequently developed the technique have been primarily interested in it as a rapid method for differential screening and no serious attempts to array the whole cDNA population of a cell type or tissue in this way have been made.

#### **1.4.3.3 The sensitivity of differential display and AP-PCR**

There are two aspects to the question of sensitivity. As mentioned earlier, both AP-PCR and differential display have been independently shown to give roughly constant patterns using the same primers with as much as 2 $\mu$ g RNA or as little as 2ng RNA (Welsh *et al.*, 1992; Liang *et al.*, 1993). Some species disappear with less RNA

and these are thought to represent false positive bands (McClelland *et al.*, 1995). The other aspect is how rare a message can be detected and how effective these techniques are at detecting differences in transcript abundance as opposed to the presence or absence of a particular species.

Bertioli and coworkers (Bertioli *et al.*, 1995) have reported a computer simulation of differential display which generated 15000 random sequences of 1200 nucleotides to represent the different mRNA species in a cell. Sequences were searched with an arbitrarily chosen 10-mer allowing up to 4 mismatches between primer and template. Their analysis suggested that a 10-mer with 4 mismatches would have an average of 20 binding sites *per* cDNA. As has already been explained not all of these would yield products, nevertheless these authors reasoned that a single 10-mer probably primes many more PCR amplicons than those visible on a differential display gel which would be amplified in the PCR until one of the substrates was depleted. Such products would remain invisible as they would not have high enough specific activity to register on autoradiography. They suggested that the dNTPs in the reaction would deplete first and at this point only those species that had reached a certain concentration would be visualised and such species would tend to be the more abundant. They had also shown that whilst differential display was unable to detect 4pg of a rare target that had been added to 2µg total RNA, specific 10-mer primers could detect this amount of target using standard reverse transcription PCR. They did not test whether other specific 10-mer primers were able to detect this target. It was suggested that sensitivity to low copy number species could be improved by the use of more efficient, *i.e.* longer, primers for PCR and by the addition of more dNTPs. AP-PCR does exhibit both of these features but whether or not this improves its detection of rare mRNAs is not known. These predictions are at variance with the practical and anecdotal experience of many workers who have observed that differential display tends to identify very low abundance mRNAs (Conway 1995; Guimarães *et al.*, 1995a,b; Rafaeloff, 1996). In addition, Liang and Pardee were able to show in their original paper that thymidylate kinase (a low copy number mRNA) could be detected using a perfectly matching 10-mer. Further experiments using 5' primers that have 2 or more mismatches to low abundance transcripts are required to fully address the issue of whether rare species can be fully represented by the method.

Ralph and coworkers have attempted to enhance the sensitivity of AP-PCR using "nested RAP" (Ralph *et al.*, 1993). Here a simple AP fingerprinting protocol is applied using the first primer, and then a small aliquot of this is reamplified using a second nested primer which is identical to the first except for the addition of one or

more arbitrarily chosen nucleotides at its 3' end. The secondary round of PCR amplifies only those molecules in the first PCR which by chance share the additional 3' nucleotides. Series of nested RAP reactions could be carried out in order to selectively sample RAP amplified cDNA for rare species that could not be seen on the original gel. This approach has not as yet been widely used.

As has been mentioned earlier, the rate of amplification of PCR products from abundant templates generally declines faster than that of rarer amplification products in the later cycles of a PCR performed on a heterogeneous template (Mathieu-Daudé *et al.*, 1996b). This occurs independently of the declining efficiency of a PCR in later cycles (which may occur on account of enzyme and substrate depletion) and is thought to be due to rehybridisation of abundant templates interfering with primer binding or extension; it has been termed the 'C<sub>0</sub>t effect'. It can result in a systematic bias against the more abundant of two products in a PCR (Innis *et al.*, 1990). Many studies in the quantitation of reverse transcription PCR, where specific primers for a known species in a cDNA population are used in PCR of cDNA and different dilutions of cloned target DNA, have predicted a C<sub>0</sub>t effect (Bouaboula *et al.*, 1992). In general, the ratios of starting template and standard are reflected in their product ratios until late in the PCR when product ratios diminish. Mathematically, the C<sub>0</sub>t effect is difficult to model as it comes into play in a PCR at the same stage as many other factors in the reaction become non-exponential and indeed it has been largely overlooked as an explanation for the decline in PCR efficiency late in the process (Mathieu-Daudé *et al.*, 1996b). It complicates the use of differential display and AP-PCR in looking for differences in abundance of a particular species between RNA populations and in the estimation of relative abundance of two species within a single RNA population as such differences would be abolished as the PCR progresses. Overall it might be predicted that differences in the intensities of identical bands in a differential display gel would underestimate the difference in the starting templates.

#### **1.4.4 Practical problems in differential display and potential solutions**

There is no doubt that reproducible band patterns can be readily generated using differential display and AP-PCR over a wide range of starting amounts of RNA from a given source (Bauer *et al.*, 1993; Liang *et al.*, 1993; Zhao *et al.*, 1996). However despite much interest in the technique as a potentially highly sensitive and rapid differential screen there were clearly problems. Most workers consistently found a very high false positive rate (up to 70%) using the method when clones were used in a secondary screen (Sompayrac *et al.*, 1995). In addition, a single differential band when cloned could give rise to multiple clones of identical size but different sequence

some of which were clearly not differential (Liang *et al.*, 1992; Welsh *et al.*, 1992; Liang *et al.*, 1993; Ayala *et al.*, 1995; Sompayrac *et al.*, 1995; Mathieu-Daudé *et al.*, 1996 a; Zhao *et al.*, 1996). Cloning artefacts (in which clones did not have either differential display primer at their insert ends) were also common. These findings are hardly surprising when one considers that the display gel shows the reaction products of a degenerate PCR carried out on a heterogeneous template. Every position on the gel may contain one or more cDNA species which may or may not be visible by autoradiography (Bertioli *et al.*, 1995). In addition, it is often technically challenging to cleanly excise a single band from a dried acrylamide gel without impinging upon areas outside that generating the autoradiographic image. The general finding in differential display of many clones that do not correspond to any sequence on the database or are undetectable by northern analysis or other secondary screen may be due to the high number of false positives cloning artefacts that are picked up.

The classical method of sifting through the clones derived from differential display or AP-PCR bands has been to sequence them individually and to assay each individual sequence represented in the band by northern or RNA protection analysis (Liang and Pardee, 1992; Bauer *et al.*, 1993; Zhao *et al.*, 1996). However, these methods require considerable amounts of RNA which may not be available to some workers, including murine embryologists. In this latter case, the small size of the clones obtained generally precludes the use of *in situ* hybridisation as an alternative secondary screen. Reverse transcription PCR using primers specific to the cloned differential PCR product has generally been of limited value as a secondary screen largely because of the cost of purchasing primers for many clones of many bands (Ariazi and Gould, 1996; Davis *et al.*, 1996).

#### **1.4.4.1 Reducing false positive bands by modifying differential display**

The standard technique for reducing false positives has been to perform the PCR reactions multiple times, often with reducing amounts of RNA (McClelland *et al.*, 1994, 1995; Mathieu-Daudé *et al.*, 1996a). It is not clear whether this is of value unless independently derived aliquots of RNA are used for the experiments. As has been mentioned, some workers have attempted to improve the reproducibility of differential display and reduce the number of false positive clones obtained by using longer primers for PCR (Ayala *et al.*, 1995; Diatchenko *et al.*, 1996a). However, the occurrence of multiple clones from one band, although claimed to be reduced, is still a problem using such strategies (Mathieu-Daudé *et al.*, 1996a). Sompayrac and co-workers (Sompayrac *et al.*, 1995) attempted to improve the method by combining the twelve possible 3' primers into three sets of four (dT<sub>12</sub>NX) with degeneracy only at

the 3' most nucleotide. This had been proposed earlier by Liang (1993) as a method for simplifying differential display. However now candidate target bands were tested by repeating the differential display PCR with each of the individual 3' primers in the set and the original arbitrary primer. The pattern of differential expression should be seen for just one of the 3' amplimers. This method will also identify but not prove to be false positive or otherwise, bands in which the arbitrary primer sequence is seen at each end. The efficacy of this is unproved. In a modification of the standard northern secondary screen, Li and coworkers have proposed that the heterogeneous differential display band is reamplified directly from the gel and used to make a mixed DNA probe for northern hybridisation. The filter blot is then cut in the area of the specific signal and the probe DNA reamplified for cloning (Li *et al.*, 1994). This method has proven to be generally problematic in the hands of many workers (Mathieu-Daudé *et al.*, 1996a)

#### **1.4.4.2 Alternative secondary screening strategies**

Despite the strategies mentioned above, false positive clones still occur and workers who have limited access to RNA still need to derive methods for economically identifying true positives. Several techniques have been proposed and all have drawbacks and inefficiencies. No single method adequately addresses the problems of the mouse embryologist.

The technique of reverse northern hybridisation in which labelled heterogeneous cDNA probes derived from the RNAs upon which the original differential display was performed are used to probe plasmid Southern blots of the clones derived from a differential display band or even replica colony lifts of these clones has been attempted (Liang *et al.*, 1995; Zhao *et al.*, 1996). Whilst rapid, the results are often ambiguous and the labelling method requires approximately 1µg RNA. An analogous method has been to probe the plasmid Southern blots with a heterogeneous probe derived from the differential display reaction itself (Vogeli-Lange *et al.*, 1996). This method however acts merely as an internal control for the differential display and does not provide relevant new information.

Many workers attempt to rapidly classify the clones they derive from a differential band by either partial conventional DNA sequencing with only one dideoxy nucleotide (“T-tracking”) or by restriction mapping of the small clones using a pair of 4-cutter restriction enzymes (Zhao *et al.*, 1996). These methods allow the researcher to rapidly ascertain the number of clones that have been generated and conventional secondary screening can be performed with each species.

Mathieu-Daudé and coworkers (Mathieu-Daudé *et al.*, 1996a) have assessed the use of single stranded conformation polymorphism (SSCP) gels as a method for rapidly purifying the product of interest from a differential display reamplified band prior to cloning. Such gels can resolve identically sized species on the basis of their sequence and consequent differential migratory properties of single stranded forms of these species in native polyacrylamide gels. In this strategy, the band of interest and the corresponding region of the negative display track are reamplified and radiolabelled by PCR. The products are heat denatured in formamide and sodium hydroxide prior to loading onto an SSCP gel. If a band was comprised of only one species, then it would give two bands on the SSCP gel corresponding to each strand. Applying the method to AP-PCR, these workers have observed that apparently single species routinely give multiple bands on SSCP gels further emphasising their heterogeneity. They assumed that products on the SSCP gel which were present in the positive track only represented the truly differential species and these could be reamplified again for cloning. Potentially, this is a way around the problem of background in differential display and AP-PCR that avoids the analysis of large numbers of clones although some form of conventional secondary screen is necessary in order to exclude a selected species as being a true false positive. Its general utility has yet to be shown.

#### **1.4.5 The technical evolution of differential display**

##### **1.4.5.1 Template preparation**

As with any endeavour of this nature, the success or otherwise of differential display is critically dependent on the quality of the starting material. Initially, Liang and Pardee used total RNA in their experiments. The RNA was treated with DNase prior to reverse transcription. Despite this measure, some reverse transcription independent bands were routinely seen and clones of amplicons derived from genomic DNA isolated (Liang *et al.*, 1993). Poly (A) selection largely removed this problem (Liang *et al.*, 1993) and the DNA fingerprint could be radically different after this procedure. The cloning of ribosomal RNA has not generally been seen when using total RNA (Zimmerman and Schultz, 1994).

##### **1.4.5.2 Primer strategies**

Many of the modifications to the primer strategies of the original protocols have already been covered. The use of longer primers either for the differential display

itself or for the reamplification steps was thought initially to improve the reproducibility of the technique and eliminate false positives (Donohue *et al.*, 1994; Ayala *et al.*, 1995; Diatchenko *et al.*, 1996a). Whilst neither of these hopes have been fully borne out (Mathieu-Daudé *et al.*, 1996a), the speed of band reamplification can be increased by their use. In their initial follow up of the clones that they isolated using differential display, Liang *et al.*, (1994) found that some reverse transcription primers were under-represented in the PCR products seen. In addition, they noted that certain two-base anchored primers gave considerable smearing in the gels. They realised that one base anchored oligo-dT primers offered excellent selectivity in subdividing the mRNA without compromising the numbers of bands seen or the readability of the gels. They were also able to append 5' restriction sites to these primers which could be of value in downstream applications.

The use of longer primers in conjunction with mixtures of thermostable polymerases *e.g.* *Taq* and *Pfu* or *Vent*, has been reported to allow the generation of fingerprints of upto 2Kb (Sokolov and Prockop, 1994; Diatchenko *et al.*, 1996a). Potentially, this could permit the production of clones long enough to make useful *in situ* hybridisation probes for secondary screening in mouse embryos. The elution of large species for reamplification is often inefficient and the specific utility of this modification in an embryological system has yet to be seen.

#### **1.4.5.3 Changes in gel technology**

The original differential display protocol used denaturing polyacrylamide “sequencing” gel electrophoresis of <sup>35</sup>S-labelled PCR products (Liang and Pardee, 1992). AP-PCR protocols have traditionally involved <sup>32</sup>P labelling (Welsh *et al.*, 1992). It was soon established that native gels were equally adequate in the resolution of PCR products and gave less complex footprints with comparable numbers of differences (Bauer *et al.*, 1993; Liang *et al.*, 1995). Eluted fragments from native gels are generally easier to reamplify, probably because of the absence of residual urea. It is widely accepted now that <sup>33</sup>P is the isotope of choice in differential display and AP-PCR (Bauer *et al.*, 1993; Liang *et al.*, 1995).

Some workers have tried to eliminate the need for any isotope in the process. “Rapid Efficient Non-radioactive” (REN) differential display has been attempted in which commercially acquired pre-cast native gels are used to resolve non-radioactive PCR products. Such gels are silver-stained in order to resolve the fingerprints (Lohmann *et al.*, 1995). This technology is costly and cumbersome as excision of differential bands is problematic. One of the first successful applications of differential display

was by Winkles and coworkers who used the technique to identify fibroblast growth factor 1 inducible genes in NIH 3T3 cells (Hsu *et al.*, 1993, 1996; Donohue *et al.*, 1994, 1995). They made numerous modifications during the process of developing the technology in their laboratory and eventually synthesised cDNA with random primers, used long arbitrary primers for PCR and displayed PCR products on agarose gels stained with ethidium bromide. Although analogous technology has been independently developed (Murphy and Pellé, 1994; Sokolov and Prockop, 1994) such an approach has met with limited success in the hands of other workers. The inherent problems of false positive clones and multiple clones from apparently single bands still pertain and it is widely held that agarose gels offer less precision than the polyacrylamide type.

#### **1.4.5.4 Fluorescent differential display using automated sequencers**

Fluorescent differential display is currently being developed in conjunction with a number of commercial organisations in order to rapidly, safely and comprehensively cover entire transcript populations and assay differences between tissue types (Ito *et al.*, 1994). Isolation of PCR amplicons is an obvious problem here and various innovative technical approaches are being developed in order to facilitate this including running PCR products in parallel for analysis with a fluorimager and excising the interesting bands from the fluorimager gel.

#### **1.4.6 Application of differential display to systems with limited RNA availability**

The vast majority of reports of novel genes cloned with both differential display and AP-PCR have come from workers in fields where it is feasible to carry out secondary screening by northern, reverse northern or RNase protection analysis for large numbers of clones (McClelland *et al.*, 1995). Where such resources are not available on account of tissue paucity, alternative screens must be sought; it has already been noted that the small sizes of the clones that are obtained using these methods largely preclude their use in the generation of riboprobes for *in situ* hybridisation analysis in many situations, including murine embryological systems. Differential display has been the technique that has been largely used in these situations.

The first published attempt to use differential display in a mouse embryological system was by Zimmerman and Schultz (Zimmerman and Schultz, 1994). They had been studying very early preimplantation embryos at the two-cell, eight-cell and blastocyst stages. They were able to generate reproducible fingerprints from the total RNA of 50 two-cell embryos. Their “secondary screen” was by sequence analysis

alone and most clones were noted to have no homologies to sequences in the public access databases available at the time. They suggested that the method could be used in conjunction with library construction from embryos derived from these stages (Rothstein *et al.*, 1992) in order to generate full length clones. These workers subsequently went on to further study one and two-cell embryos using differential display and used reverse transcription PCR analysis with differential display clone specific primers in their secondary screen (Davis *et al.*, 1996). They were able to show that a translation initiation factor was up-regulated during the two-cell stage of development in both of their screening steps.

A more ambitious study was performed to compare RNAs derived from dissected 8.5 *dpc* embryonic yolk sacs, whole embryos at 8.5 *dpc*, crude dissections of 8.0 *dpc* mouse embryos corresponding to the head primordia and posterior body regions respectively, ES cells and embryoid bodies (Guimarães *et al.*, 1995a). These workers were interested in haematopoietic stem cells and using differential display to compare RNA's from tissues with different haematopoietic activities. Secondary screening by Northern blot analysis was possible in this situation and a novel vertebrate nucleobase permease cloned from the yolk sac. No *in situ* analysis was attempted. More success has been achieved in the application of differential display to lower vertebrate embryological systems where tissue and RNA are more abundant and conventional secondary screening by northern blot can be employed. In a study comparing blastula and gastrula poly (A) enriched RNA from the zebrafish, Conway (1995) isolated a gastrula specific cDNA amplicon which he could verify by northern. The clone was unsuitable for *in situ* hybridisation analysis and a full length species was isolated for this purpose from a conventional library. He was able to show that for this species at least, the temporal patterns of expression seen by *in situ* were mimicked by reverse transcription PCR analysis of RNA's from different stages of development. A false positive rate of greater than 60% was seen. Similar studies have been carried out in late gastrula stage *Xenopus* embryos that had been treated with retinoic acid and compared with untreated embryos (Mason *et al.*, 1996). Here the secondary screen was by RNase protection. *In situ* hybridisation analysis was only attempted after a full length clone was isolated from a library.

There has yet to be a report of a screening experiment involving an arbitrarily primed PCR method in which dissected mouse embryonic tissues have been compared; in such an experiment, *in situ* hybridisation studies would have to comprise the secondary screen. Encouragingly, there has been one recent report of the successful use of *in situ* hybridisation as a secondary screen in a differential display experiment. In a study comparing RNA derived from punch skin biopsies of normal volunteers

and the active lesions of psoriatic patients, using a standard PCR protocol, Rivas and colleagues claimed that eight out of twenty clones analysed were verified by *in situ* hybridisation as representing species enriched in psoriatic skin (Rivas *et al.*, 1997). They used radioactively labelled riboprobes and frozen sections of the clinical material. The *in situ* labelled sections often showed considerable background hybridisation to the naked eye and so computer software was employed to quantify the density of silver grains in different anatomical regions. The authors commented that small clones generally gave poor probes yielding no result in the secondary screen and so were discarded. No information was presented regarding the percentage of bands that were not studied on account of this.

#### **1.4.6.1 Differential display at the level of a single cell: *in situ* transcription**

The technique of *in situ* transcription was initially developed in order to synthesise cDNA from low abundance cells whose precise anatomical location is known in a tissue or organ but for which large scale enrichment or dissection was not possible (Tecott *et al.*, 1988). Gene specific oligonucleotide primers were annealed to the mRNA in a tissue section and reverse transcription initiated *in situ* by the addition of enzyme and substrates, allowing the incorporation of radiolabelled deoxynucleotides. Autoradiographic exposure of the tissues under study would then give an idea of cellular localisation of specific transcripts. The method could be readily adapted for use with mouse embryo sections and cDNA eluted for the purposes of PCR amplification and cloning (Eberwine *et al.*, 1992).

Similar *in situ* transcription studies could be performed in single live neurones (which could be identified and isolated in low density cultures of isolated brain tissues) by microinjection of cDNA synthesis reagents into the cell through patch-clamp recording pipettes, with subsequent aspiration of the entire cellular contents (Eberwine *et al.*, 1991). Here, the reverse transcription primer was oligo d(T<sub>24</sub>) with a bacterial T7 RNA polymerase promoter sequence appended to its 5' end and cDNA synthesis was actually carried out in the micropipette. After conventional second strand synthesis the cDNA could be amplified using T7 RNA polymerase to give antisense RNA (aRNA), a single round of RNA synthesis yielding a 2000-fold RNA amplification. The aRNA itself could then be cloned in the form of a cDNA library or used to differentially screen pre-made cDNA libraries encompassing larger masses of tissue.

In subsequent studies involving *in situ* transcription, the same authors were able to use aRNA from isolated hippocampal cells for reverse northern analysis of known

clones and conventional differential display (Mackler and Eberwine, 1994; Miyashiro *et al.*, 1994). In principle it is possible to use aRNA as a template for a secondary screen of differential display amplicons though this has yet to be reported. Potentially the technology could be of value in murine embryology, though the patch-clamp technique will be unfamiliar in many laboratories.

### **1.5. Other novel methods of screening for differential gene expression.**

All of the methods mentioned thus far have the major drawback of a high false positive rate. In the case of differential display, the fact that the clones isolated are merely fragments of candidate genes often corresponding to the 3' UTR and not full length species further increases the potential work involved in the screen. In efforts to either refine existing strategies or to generate entirely new ways of differential screening, many novel and interesting techniques have been tried some of which may appear to be of use to the mouse embryologist. Unfortunately problems exist with them all and no single technique has come to the fore.

#### **1.5.1 Domain specific differential display**

This involves reverse transcription with a short arbitrary primer followed by PCR with that primer and a degenerate primer homologous to a conserved region within a gene family *e.g.* zinc finger domains (Stone and Wharton, 1994). AP-PCR parameters are generally employed. Some workers have carried out analogous studies using conventional differential display reverse transcription primers (Johnson *et al.*, 1996). The process can be coupled with restriction analysis of the PCR products in order to further distinguish different members of the same family of genes which may give identically sized products (Fischer *et al.*, 1995). This approach can be viewed as a variant of PCR experiments with degenerate primers used to “fish” for genes with specific motifs except that only one specific primer is required. Differential clones can be sequenced and used with a degree of confidence in library screening.

##### **1.5.1.1 Multiplex profiling**

Multiplex profiling involves the reverse transcription of random primed cDNA with a mixture of several gene or domain specific primers in a single tube (Jensen *et al.*, 1996). As with domain specific differential display, an arbitrary primer is usually employed in combination. The reaction products can be visualised after electrophoresis and differential species cloned, used for blotting and probed with

gene family specific probes or used in further differential display experiments as a form of enriched template.

### **1.5.2 Fingerprinting of 3' end cDNA fragments**

Two strategies have been independently developed which aim to differentially fingerprint the 3' ends of cDNAs using a combination of restriction analysis and PCR. As in differential display, both methods attempt to reduce the complexity of the fingerprints obtained by either reverse transcription of only a fraction of all messages with anchored oligo dT primers or by prior restriction digestion of total cDNA with one of a panel of enzymes. However both have the potential advantage over differential display of giving non-redundant information.

#### **1.5.2.1 Differential 3' end cDNA profiling**

This strategy was developed with a view to reducing the high false positive rates of differential display (Prashar and Weissman, 1996). A panel of two base anchored dT<sub>18</sub> primers is used which have long (20 base) 5' extensions to prime the synthesis of double stranded cDNAs from the RNAs under study which are in turn digested with a restriction enzyme selected from a panel of 6-cutters. The cut cDNA is ligated to an adapter that has a 3' overhang for ligation and noncomplementary sequences at its 5' ends, thus giving it a Y-shape. The entire ligated cDNA is then PCR amplified at high stringency using a radiolabelled primer complementary to the 5' extension of the reverse transcription primer and an unlabelled primer complementary to the 5' end of the adapter. Only species which have the reverse transcription primer at their end will be seen after electrophoresis. Differentially expressed species can be reamplified and cloned as in differential display and AP-PCR. The proponents of this technique claim less than 5% false positives and near total representation of all mRNAs using the combinations of reverse transcription primers and 6-cutter enzymes (S.Weissman Yale Medical School, personal communication). The published method requires at least 10µg total RNA.

#### **1.5.2.2 Gene expression fingerprinting (GEF)**

In this similar method, conventional cDNA synthesis is performed with approximately 5µg of each RNA using a single biotinylated oligo dT primer and digested with a frequently cutting restriction enzyme. 3' end directed restriction fragments are separated with streptavidin and ligated with an adapter to their restricted 5' ends. A PCR is then performed using an adapter primer and the

biotinylated oligo dT primer and after immobilisation of the products with streptavidin the amplified cDNA is radiolabelled using the adapter primer and T7 polymerase. The immobilised radiolabelled cDNA can then be digested sequentially with a panel of restriction enzymes and released fragments electrophoresed in differential display gels (Ivanova and Belyavsky, 1995). Fingerprints for as many enzymes as possible can be made and the authors suggest that this will allow near total representation of all mRNAs. Differentially expressed species that are eluted from the gels are PCR amplified after homopolymer tailing with terminal transferase.

For both of these techniques large amounts of RNA were required in the original protocols. It may however be possible to modify them for analysis of PCR amplified total cDNA from dissected embryonic structures.

### **1.5.3 Combining subtraction and differential display**

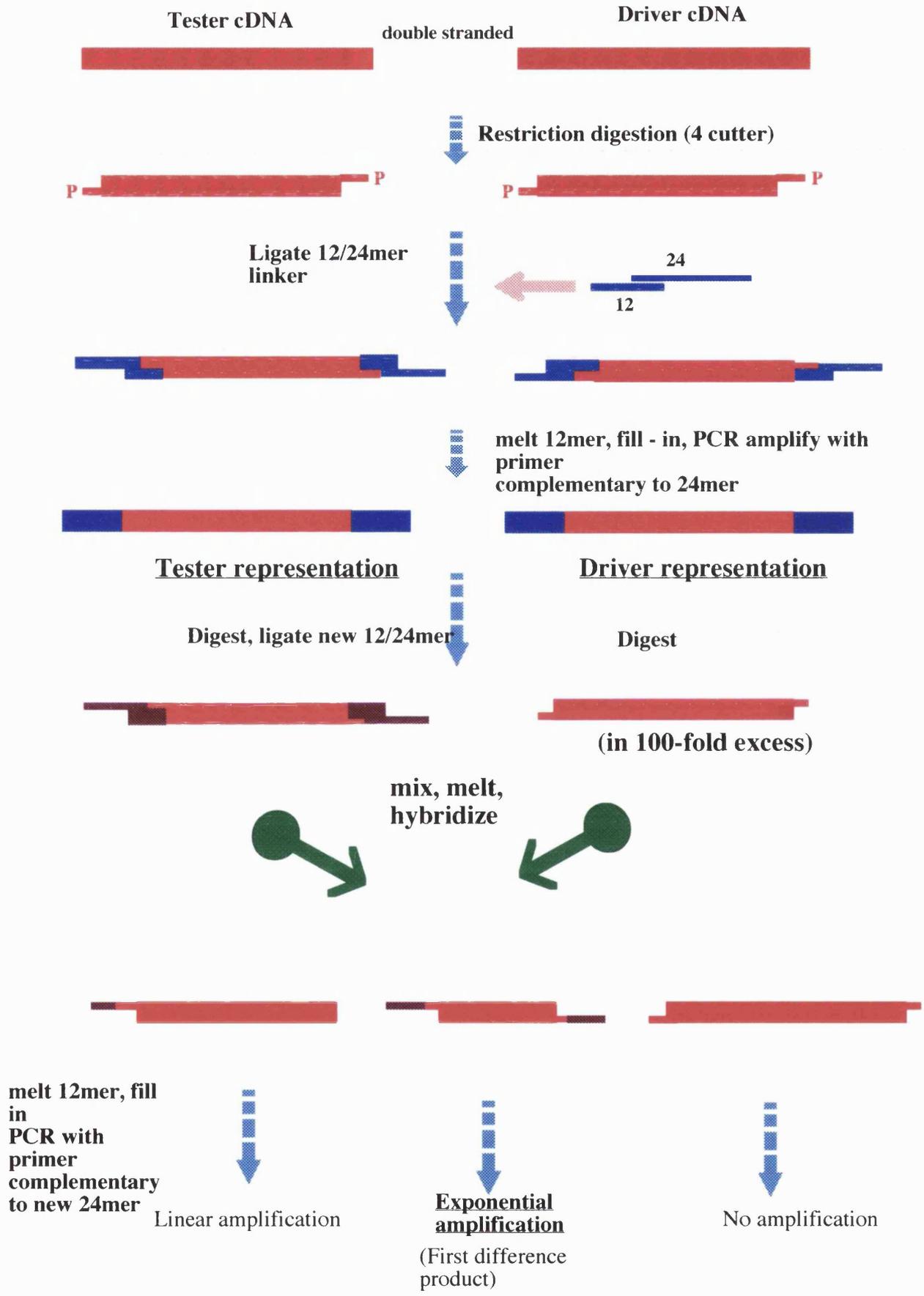
The term “subtractive display” has been applied to differential display performed on subtracted cDNA (Hakvoort *et al.*, 1994; Ariazi and Gould, 1996). The total cDNAs to be compared are tailed and PCR amplified prior to subtraction. The method of Ariazi and Gould entails cloning a fraction of the cDNA after each round of subtraction and PCR fingerprinting of each cloned aliquot with an arbitrary primer and a vector primer. The value of the subtraction step *per se* is not clear but the principle of differential display of PCR amplified cDNA is intriguing for the murine embryologist.

#### **1.5.3.1 Representational difference analysis (RDA)**

This is a PCR-coupled subtractive process based on a method for DNA difference enrichment that was initially developed in order to clone the differences between complex genomes (Lisitsyn *et al.*, 1993; Hubank and Schatz, 1994) (figure 1.2). The aim is to deplete the RNA populations being compared of common sequences and PCR amplify the differences. Double stranded cDNA is prepared from tester and driver RNA sources, digested with a 4-cutter enzyme and ligated with a nonphosphorylated 12/24 adapter composed of one strand of 24 bases and the other of 12 bases. Only the 24-mer strand is covalently attached to the 5' ends of the cDNA and so the 12-mer species is melted off and the cDNAs are amplified with a primer complementary to the 24-mer. The adapter sequences are then removed from the cDNAs by restriction digestion and the amplified tester cDNA ligated with a different 12/24 adapter before a subtractive hybridisation is performed with a 100-fold excess of driver to tester. After the hybridisation the mixture is PCR amplified

### **Figure 1.2**

Schematic representation of RDA (after Hubank and Schatz(1994)). Amplified tester and driver cDNA populations are derived by PCR as described in the text. After restriction digestion, a 12/24mer linker is ligated (shown in blue). As the linker is nonphosphorylated, only the 24mer component is ligated to digested cDNA. The cDNA is digested with a second enzyme and a second linker (shown in brown) ligated to the tester cDNA only. After hybridisation, the only cDNAs which can be amplified in a second round of PCR using primers complementary to the second linker are those which are enriched in the tester population. The resulting first difference product may be enriched by further rounds of hybridisation and PCR.



**Figure 1.2**

with a primer complementary to the 24-mer of the second adapter. Only tester/tester homohybrid sequences should be amplified as double stranded DNA and single stranded DNA derived from the amplification of tester/driver heterohybrids can be removed enzymatically.

After several rounds of subtraction, enriched tester specific species can be viewed after agarose gel electrophoresis and cloned. The method requires large amounts of RNA.

### **1.5.3.2 Suppression subtractive hybridisation (SSH)**

This method is similar to RDA and was initially developed as a PCR based approach for the generation of subtracted cDNA for library construction (Diatchenko *et al.*, 1996b). Tester and driver RNAs are converted to double stranded cDNAs using conventional methods. Both are then digested with a 4-cutter enzyme that yields blunt ends; the tester population is split into two samples and each ligated with a different nonphosphorylated adapter. An excess of driver is then added to each tester sample and the cDNAs are heat denatured and then allowed to reanneal. Each cDNA mixture should now be enriched for single stranded differentially expressed genes (Figure 1.3). If the two are now mixed and further competed with more heat denatured driver, then a fraction of the single stranded species generated after the first subtractions can reanneal such that they have one of each adapter sequence at their 5' ends. These species can be specifically PCR amplified with primers complementary to the 5' halves of each adapter and the radiolabelled reaction products visualised after polyacrylamide gel electrophoresis (figure 1.3).

This somewhat complicated technique also requires large amounts of RNA and has yet to find wide usage.

### **1.5.4 Serial analysis of gene expression (SAGE)**

This interesting method allows for the rapid simultaneous analysis of many short (9 - 13bp) sequence tags which represent a "snapshot" of the genes expressed in a particular tissue (Velculescu *et al.*, 1995). Although not strictly a method of screening for differential gene expression, it permits the speedy identification of sequences that might represent tissue specific transcripts.

The basic idea is that a short sequence tag of a few base pairs contains enough information to uniquely identify a transcript provided it is isolated from a defined

### **Figure 1.3**

Schematic representation of SSH (after Diatchenko *et al.* (1996b)). Tester cDNA is split into two halves by digestion with a 4-cutter restriction enzyme and each is ligated to a different nonphosphorylated adapter bearing an inverted repeat sequence (indicated in red and pink). Adapters can only ligate to the phosphorylated 5' ends in each cDNA population. Each tester pool is then hybridised with an excess of driver cDNA and the four possible classes of reannealing products indicated (A, B, C and D) are produced. The tester pools are then mixed and reannealed with more excess driver cDNA. The reannealing products produced by this second hybridisation are shown. If the ends of these cDNA molecules are then filled-in and PCR amplified with nested primers complementary to part of each adapter, then the only species that will be amplified are those present in each tester population which have different adapter sequences at each end. In principle, these should be enriched in the tester population. Species of class C may also be enriched in the tester cDNA but these will not be amplified on account of the inverted repeat in the adapter which favours self priming in the PCR. The amplified cDNA products can be cloned or used for another round of enrichment.

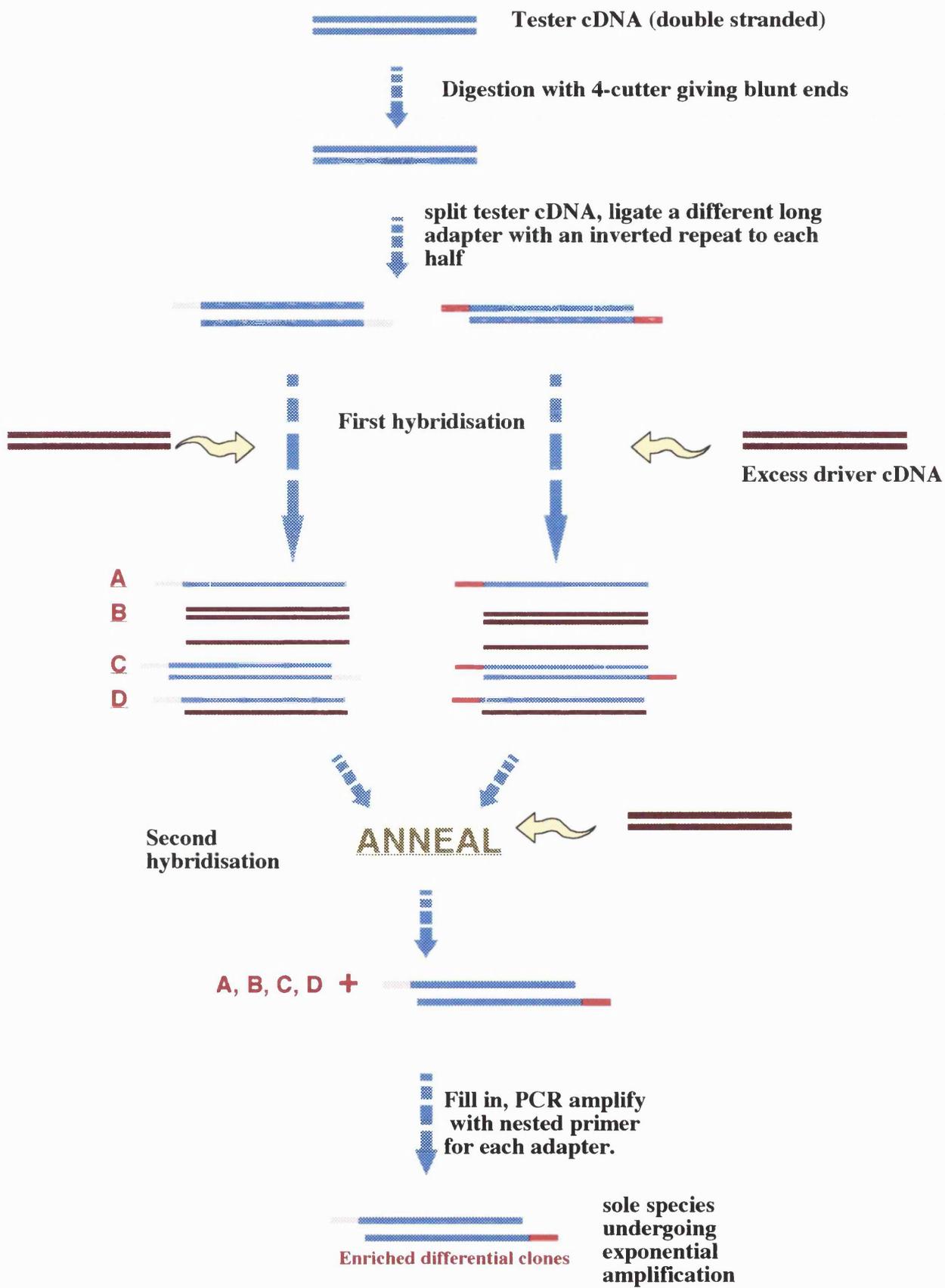
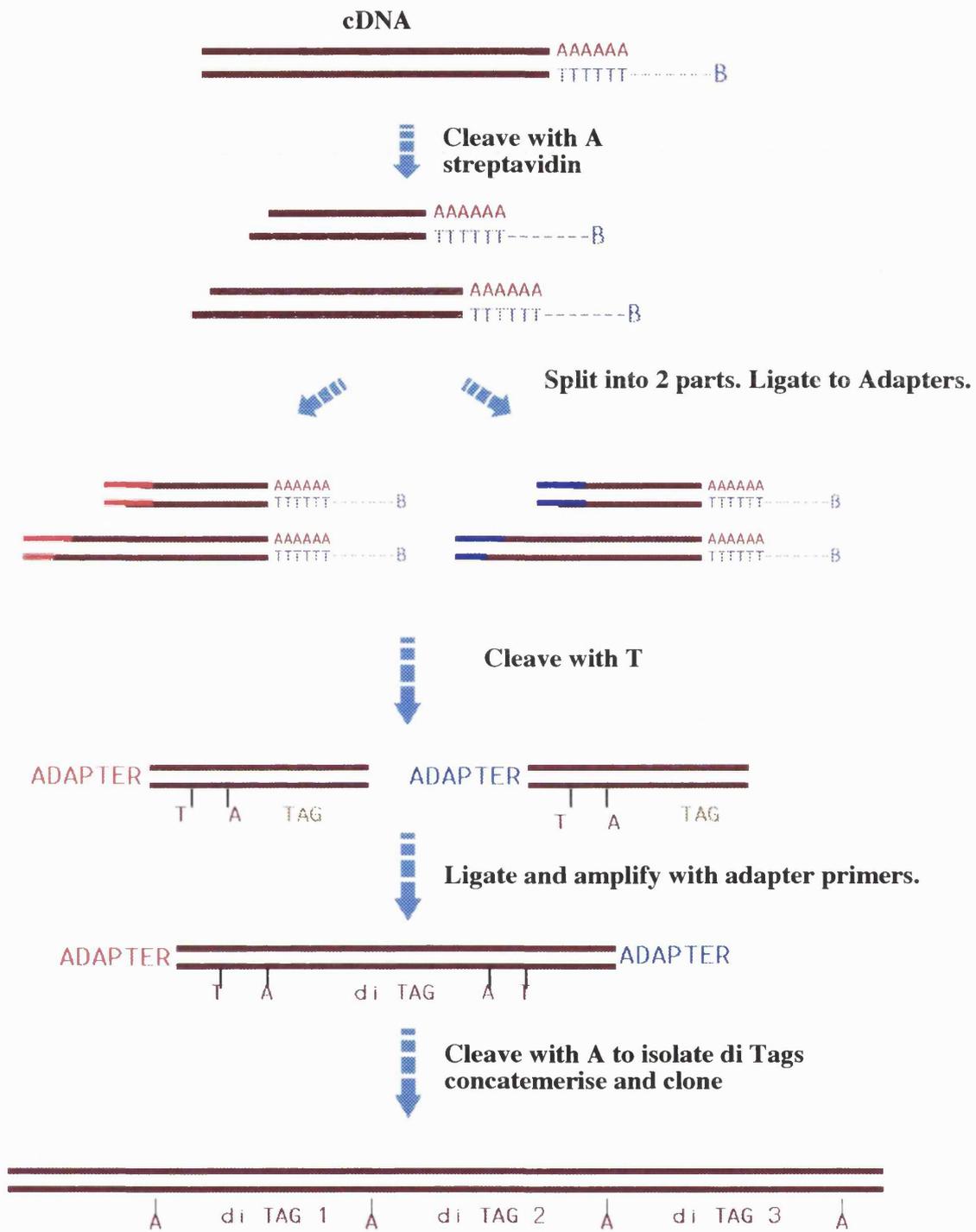


Figure 1.3

#### **Figure 1.4**

Schematic outline of the SAGE method (based on Velculescu *et al.* (1995)). Double stranded cDNA is synthesised conventionally after reverse transcription with a biotinylated dT primer (indicated by B) and cleaved with a frequent cutting "anchoring" restriction enzyme (A). After purification with streptavidin the 3' directed cDNA digestion products are split into two halves and each ligated to a different double stranded adapter (indicated in blue or red) bearing the recognition site for a class IIS "tagging" enzyme (T). This leaves short products with 9 - 13bp sequence tags representing the cDNA population. These are ligated and amplified with primers specific to each adapter to give amplified dimerised tags which can be released after digestion with A, concatemerised and cloned. Clones can be size selected to contain an array of tags representing expressed sequences in the cDNA population.



**Figure 1.4**

part of it *e.g.* the 3' untranslated region. In the method, double stranded cDNA is synthesised from the RNA of interest using biotinylated oligo - dT (figure 1.4). This is cleaved with an “anchoring” enzyme typically a 4 - cutter that would be expected to cut in most transcripts. The resulting 3' directed cDNA is purified with streptavidin beads and divided into two equal halves and each half ligated to a different linker via the anchoring restriction site. Each linker contains a “tagging” enzyme restriction site of the type IIS class. This class of restriction enzyme *e.g.* *BsmF1* and *FokI* cleave at defined distances upto 20bp away from their asymmetric recognition sites. After digestion of the linkered cDNA populations with the tagging enzymes, the two pools of released tags are ligated and PCR amplified with primers specific to each linker. The resulting amplicons should all contain two different tags (“ditags”) and these can be cleaved using the anchoring enzyme, concatenated and cloned.

The authors claim that if these cloned concatemers are sequenced, then information regarding the corresponding sequence abundance of the cDNAs from which they derive can be obtained. If enough tagging enzymes are used then potentially the entire mRNA array of a particular tissue could be displayed. It is claimed that the short sequence tags that are obtained are sufficient for database interrogation and for library screening. For the most biologists however, they are essentially useless and the method has not been widely applied in pure research. It requires microgram quantities of RNA and the tagging enzymes often do not cut properly (Dr. D. Lockhart, Affymetrix, Santa Clara USA, personal communication). Although some commercial interest in the method has been shown, the idea is being superseded by newer technologies for the arraying of cDNAs (*vide infra*).

### **1.5.5 Expression monitoring by hybridisation to high density cDNA arrays**

Technology has recently been developed that allows for the generation of arrays of thousands of chemically synthesised oligonucleotide probes on small polymer “chips” of approximately 2cm square (Chee *et al.*, 1996). The sequences on any given chip can be custom made to correspond to any gene family or domain and can be multiply probed with fluorescently labelled antisense RNA made from as little as 100ng poly (A)<sup>+</sup> RNA as described earlier (Eberwine *et al.*, 1992). Probes can be arranged such that a probe that exactly matches a known motif sequence is placed adjacent to every possible one, two or three base pair mismatch for that sequence.

The oligonucleotide arrays are “read” after hybridisation using a computerised chip reader and the inventors claim a sensitivity of 1 copy per cell with a dynamic range of up to 10000 copies per cell in the same experiment. The oligonucleotide sequence

corresponding to a positive signal can be accessed in under 15 minutes from the start of a scan. They can be used to rapidly and directly access the growing body of cDNA sequence data without the need to analyse thousands of cDNAs.

Currently single chips are available that contain 16000 20-mers corresponding to homeodomain sequences and new versions are being produced which cover all available human EST sequences. In the future this may prove to be the method of choice for differential screening if random oligonucleotide chips are generated.

### **1.6. Systematic comparisons of techniques for differential screening**

There have been no published studies that have compared the relative efficiencies of the library based or PCR-based differential screening methodologies. If the isolated report of Bettenhausen and Gossler, (1995) is to be believed then these techniques are superfluous and all that is necessary is to pick random clones from a library of interest. Sadly this has generally been an ineffective technique in the hands of others.

At the recent Cold Spring Harbor Laboratory symposium on differential display, Dr. Mark Erlander of the R.W. Johnson Pharmaceutical Research Institute in San Diego reported the preliminary results of a formal comparison of differential display and library subtraction in the search for interferon  $\gamma$  inducible genes in HeLa cells. Using Northern analysis as a secondary screen, he found that of 186 cDNAs selected after library subtraction 33 were true positives and of 80 differential display bands analysed 39 were positive. Though these data suggest that differential display is more efficient in his hands, the pooled results of many more studies like this in which identical experimental techniques are used would have to be analysed before a definitive conclusion could be reached.

At the moment the investigator has to choose between the relative merits and drawbacks of each method and use his or her personal preferences in deciding the technique to use. For the mouse embryologist, the choice is further limited by the issues of RNA availability and an appropriate secondary screen.

### **1.7. Somite formation in the mouse embryo**

Segmentation is a fundamental process in vertebrate embryology and occurs after the completion of germ layer formation. It allows for the subsequent spatial organisation of determined cells and differentiating tissues in the developing embryo (Tam and

Trainor, 1994). In the mouse, the first morphological signs of segmentation are the formation of neuromeres in the neural tube and of somites in the mesoderm.

The tissues of the mouse embryo are derived from a single epithelial sheet of cells shaped like a hollow cone, the epiblast (Tam and Beddington, 1987). During the process of gastrulation, this structure generates the three germ layers; an inner layer of ectoderm and an outer layer of endoderm with an intervening layer of mesoderm. Each of these layers will give rise to specific parts of the developing fetus. The mesoderm and endoderm arise from the epiblast at the site of a specialised group of cells which run down one of its sides. This structure is the primitive streak and it is thought that the allocation of cells to different mesodermal lineages may be related to the sites at which they ingress through the streak (Tam and Beddington, 1987). A schematic of this is shown in figure 1.5a.

There are several anatomical subsets of mesoderm and it is the paraxial tissue in which somites derive, the newest somites lying most caudally. As the name suggests, the paraxial mesoderm lies along the longest axis of the embryo. Orthotopic grafting studies in mice suggest that it arises from the middle portion of the primitive streak (Tam and Beddington, 1987). Scanning electron microscopy of the outwardly unsegmented paraxial mesoderm caudal to the newest and rostral to the oldest somite in avian embryos has revealed series of spherical clusters of cells called somitomeres (Tam and Trainor, 1994). Whilst somitomeres in the cranial mesoderm never form somites, those located caudal to the newest somite (in the presomitic mesoderm) will segment into full formed somites (figure 1.5b).

The somitic pattern along the rostro-caudal axis of the embryo provides a basis for the generation of the vertebral column and the muscle groups at different levels of the body. The ventral part of each somite subsequently forms sclerotome, which gives rise to the axial skeleton and the dorsal portion forms dermomyotome, precursor tissue for dermis and skeletal muscle. Each dermomyotome can be further divided into compartments giving rise to axial and limb/body wall musculature. The formation of these somite lineages occurs at constant times after segmentation in each species and is thought to be regulated by complex patterns of signals impinging on the somite from adjacent structures specific to that somite's position in the embryo. Each somite contributes to structures that are always within one body segment, suggesting that the segmental pattern of the somites is maintained during subsequent development. When individual chick somites are surgically removed, adjacent somites never compensate the loss (Bagnall *et al.*, 1988).

### **Figure 1.5**

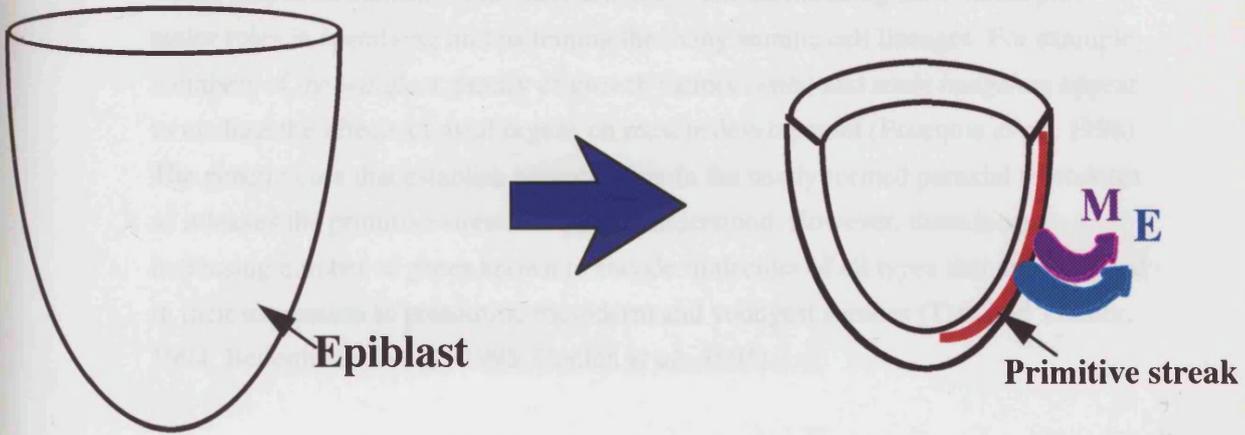
(a) Schematic representation of the formation of mesoderm and endoderm germ layers by the epiblast at gastrulation. A cross-section of the epiblast that exposes the primitive streak at the posterior side of the embryo is shown. The emergence of cells destined to be mesoderm (m) and endoderm (e) is depicted by the arrows.

(b) Schematic representation of somitogenesis in the paraxial mesoderm (after Tam and Trainor, 1994, figure 3). Cells are recruited into the caudal end of the paraxial mesoderm after ingress from the primitive streak. They are then organised into somitomeres which mature by increasing their cell density. Segmentation is a progression of this and is heralded by the compaction of cells and the formation of intersomitic clefts. The first lineage differentiation is marked by the dissociation of epithelial cells on the ventro-medial side of the somite to form the sclerotome.

(a)

position of  
extra-  
embryonic  
structures

Cross-section



Anterior

Posterior



(b)

Somites

Presomitic mesoderm

8 7 6 5 4 3 2 1

dermamyotome

Anterior primitive  
streak



sclerotome

Tail bud

lineage  
differentiation

segmentation  
of somites

organisation  
into  
somitomers

Recruitment of  
paraxial mesoderm

Figure 1.5

Nascent somites as they bud off, appear as paired epithelial spheres on either side of the midline. As the subdivisions of the somite form, it seems that the cells from which they arise condense into mesenchyme. Cells surrounding the somites play major roles in specifying and patterning the many somitic cell lineages. For example, members of the *wingless* family of growth factors (*wnts*) and *sonic hedgehog* appear to mediate the effects of axial organs on muscle development (Pourquie *et al.*, 1996). The genetic cues that establish segmentation in the newly formed paraxial mesoderm as it leaves the primitive streak are poorly understood. However, there is an increasing number of genes known to encode molecules of all types that are restricted in their expression to presomitic mesoderm and youngest somites (Tam and Trainor, 1994; Bettenhausen *et al.*, 1995; Conlon *et al.*, 1995).

### **1.7.1 The role of cell surface molecules in somitogenesis**

One of the fundamental findings in molecular embryology has been that many genes involved in important developmental events in *Drosophila* are structurally and functionally conserved at some level in vertebrates. The *notch* locus in *Drosophila* encodes a transmembrane molecule essential for the correct epidermal versus neural cell fate decisions in the neurogenic region of the fly ectoderm (Greenwald and Rubin, 1992). Homozygosity for null mutations at the *notch* locus is lethal in *Drosophila* embryos - essentially all cells in the neurogenic region become neuroblasts, and such embryos die with hypertrophy of the central nervous system and absence of epidermal structures. The known ligands for *notch* in the fly are the membrane bound *delta* and *serrate* gene products, while *suppressor of hairless* and *deltex* gene products are involved in intracellular *notch* signalling. There is a large body of experimental data which suggests that *notch* functions in inductive signalling, cellular pathfinding and lateral inhibition (Greenwald and Rubin, 1992; Artavanis - Tsakonas *et al.*, 1995). This latter phenomenon refers to the differentiation of certain cells within an initially equivalent population to a particular fate whereas inductive signalling refers to two non-equivalent cell populations, one instructing the other to a particular fate. This signalling system, which is still incompletely understood, is also utilised at other developmental stages in the fly.

There are at least four different *Notch* homologues in vertebrates and several homologues of *delta* and *serrate* and their expression patterns are dynamic and complex during early nervous system formation, gastrulation and somitogenesis (Williams *et al.*, 1995; Bettenhausen *et al.*, 1995). *In situ* hybridisation studies of the expression of *notch* homologues in early mouse embryos have shown that all three

are expressed at some stage in somitogenesis. *Notch 1* is highly expressed in presomitic mesoderm and nascent somites after which it declines to the dorsomedial region of the somite; *notch 2* is expressed only after the formation of somites and progressively diminishes and finally, *notch 3* is seen in the somites only (Williams *et al.*, 1995). Targeted mutagenesis studies of the *notch 1* locus in mice have shown that mouse embryos homozygous for non-functional *notch 1* alleles are developmentally retarded and die by 11.5 *dpc* with somite condensation coming to a halt after the formation of the first 14-16 somites (Swiatek *et al.*, 1994). Localised areas of cell death are observed in the developing nervous system in such animals.

### **1.7.2 The role of transcription factors in somitogenesis**

Using a yeast GAL4 two-hybrid system to screen a murine 14.5 *dpc* embryo fusion expression cDNA library for E12 dimerisation partners, Cserjesi *et al.* (1995) were able to identify a novel basic helix-loop-helix (bHLH) transcription factor gene, *scleraxis*. Transcripts for this gene were found from the onset of gastrulation but expression was marked in the lateral sclerotome and mesenchyme of the limb buds and body wall from 10 *dpc*. Subsequently more widespread expression could be observed throughout the axial skeleton.

The same group then attempted to isolate related genes by low stringency screening of a murine 13.5 *dpc* embryo cDNA library with a probe specific for the bHLH domain of *scleraxis* (Burgess *et al.*, 1995). A related gene, *paraxis*, was cloned which was observed on *in situ* hybridisation analysis to be first expressed at 7.5 *dpc* in posterior primitive mesoderm, subsequently appearing in the most rostral paraxial mesoderm (from which somites are about to form) and then in all three compartments of newly formed somites. As the somites mature and move rostrally, expression declines and by 13.5 *dpc*, is undetectable. Targeted disruption of the *paraxis* gene led to early neonatal death in mice with obvious gross skeletal dysgenesis (Burgess *et al.*, 1996). Early embryos homozygous for the mutation failed to show complete somitogenesis, instead the paraxial mesoderm was segmented into loosely defined mesenchymal units of the approximate size and periodicity of somites, suggesting a functional separation of somitogenesis and segmentation of the paraxial mesoderm and a role for *paraxis* in the former process. Interestingly, *in situ* analyses of these embryos also revealed that they expressed dermatomal, myotomal and sclerotomal markers in their loosely condensed presomitic mesoderm though their boundaries of expression were less well defined than wild type. Thus there is molecular specification of somitic cell lineages and domains despite the failure of somite

condensation and underlying the somites, there exists an obscure segmental organisation of the paraxial mesoderm.

Despite recent advances, the picture of how somitogenesis comes about is still patchy and the full repertoire of genes involved is far from complete. No differential cDNA screening study searching for differences in gene expression between murine somites at different axial levels or comparing somitic and presomitic paraxial mesoderm has been reported for the reasons discussed earlier. If the problems of secondary screening and a high false positive rate could be overcome, differential display may be a suitable technique for attempting such a project.

### **1.8. Aim**

The work presented subsequently will assess the suitability of differential display as a method of screening for differential gene expression between dissected somites and presomitic mesoderm from early mouse embryos. It will follow the development of the technology geared towards working with small amounts of tissues, and will describe the technical modifications necessary to reduce background and allow for secondary screening by *in situ* hybridisation.

## CHAPTER 2

### Materials and methods

#### 2.1 MATERIALS

##### 2.1.1 Bacterial host cell strains

**Epicurian Coli™ NM522** (Stratagene)

*supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5 (r<sub>k</sub>, m<sub>k</sub>) [F' proAB, lacI<sup>q</sup>ZΔM15]*

**Epicurian Coli™ XL1-Blue**(Stratagene)

*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI<sup>q</sup>ZΔM15*

*Tn10(tet<sup>r</sup>)]*

**Epicurian Coli™ XL1-Blue MRF'**(Stratagene)

*Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr) recA1 endA1 gyrA96 thi-1 hsdR17 supE44, relA1 lac [F' proAB, lacI<sup>q</sup>ZΔM15 Tn10(tet<sup>r</sup>)]*

**Escherichia Coli SOLR e14<sup>-</sup>** (McrA<sup>-</sup>) *Δ(mcrCB-hsdSMR-mrr)171 sbcC recB recJ uvrC umuC::Tn5 (Kan<sup>r</sup>) lac recA1 endA1 gyrA96 thi-1 λ<sup>R</sup> [F' proAB, lacI<sup>q</sup>ZΔM15]<sup>c</sup>*  
(non - suppressing)

##### 2.1.2 General chemical reagents and solutions

All reagents were prepared using distilled, deionised water. Solutions used in the manipulation of RNA were further treated with 0.1% (v/v) diethyl pyrocarbonate as described below (Sambrook *et al.*, 1989). NIMR laboratory safety guidelines and manufacturers' warnings were heeded at all times.

**Acetone** (BDH)

**38% acrylamide/2% bisacrylamide (w/v) stock** (Amresco)

**Agarose** (DNA grade, Biorad).

**Agarose gel loading buffer (DNA)** 20% (w/v) ficoll 400, 0.1M EDTA, 0.25% (w/v) bromophenol blue (Sigma), 0.25% (w/v) xylene cyanol (Sigma).

**Alkaline agarose gel loading buffer (DNA)** 20% (w/v) ficoll 400, 50mM NaOH, 2.5 mM EDTA, 0.25% (w/v) bromophenol blue (Sigma).

**Alkaline agarose gel running buffer** 50mM NaOH, 2.5mM EDTA

**Ammonium persulphate** 10% (w/v) aqueous solution stored at 4°C.

**Ampicillin** (Sigma) Stock aqueous solution 100mg/ml (filter sterilised). Working concentrations 50mg/ml (plates) and 25mg/ml (broth).

**Anti - digoxigenin Fab fragments** (Boehringer Mannheim). 0.5µg/ml in buffer I. Stored at 4°C.

[ $\alpha$ <sup>33</sup>P]dATP (1000-3000 Ci/mmol, Amersham).

[ $\alpha$ <sup>35</sup>S]dATP (400Ci/mmol, Amersham).

**BCIP** 5-Bromo-4-chloro-3-indolyl-phosphate, 4-toluidine salt (Boehringer Mannheim). 50mg/ml solution in dimethylformamide. Stored at -20°C away from light.

**BHI medium** (Brain Heart Infusion broth) Difco BHI powder 37g/l. Sterilised by autoclaving for 20 minutes at 121°C, 15lb/in<sup>2</sup>.

**Blocking Powder** (Boehringer Mannheim)

**BSA (Bovine serum albumin)** (Sigma).

**Buffer I (Whole mount)** 0.15M NaCl, 0.1M Tris-HCl pH 9.5

**Buffer II (Whole mount)** 1% (w/v) Blocking reagent (Boehringer Mannheim) in buffer I

**Buffer III (Whole mount)** 0.1M NaCl, 0.1M Tris-HCl pH 9.5, 50mM MgCl<sub>2</sub>

**Butan-1-ol** (Merck). Equilibrated with water by mixing in a 2:1 ratio at room temperature.

**CHAPS** (Sigma). 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulphonate.

**Chloroform** (Rectapur<sup>TM</sup>, Prolabo).

**Chloroform/iso-amyl alcohol** 49/1 mixture prepared as required.

**Column wash buffer (riboprobe)** 10mM Tris-HCl pH 8, 1mM EDTA, 10 mM DTT and 0.1µg/µl yeast tRNA

[ $\alpha^{32}\text{P}$ ]dCTP (1000-3000 Ci/mmol Amersham).

**Deionised formamide** 500ml of fresh formamide (BDH) were mixed with 10mg mixed bed ion-exchange resin (Bio-Rad AG<sup>®</sup> 501 - X8, 20 to 50 mesh) and stirred for 30 minutes at room temperature prior to filter sterilisation. Stored at -20°C.

**Denaturing solution (RNA extraction)** 4M guanidinium thiocyanate, 25mM sodium citrate pH 7, 0.5% (w/v) sarkosyl, 0.1M 2-mercaptoethanol. Stored at -70°C.

**Denhardt's solution (100x)** 2% (w/v) BSA, 2% (w/v) ficoll 400, 2% (w/v) polyvinylpyrrolidone, filter sterilised and stored at -20°C.

**Diethyl pyrocarbonate (DEPC) treatment of Solutions** 0.2 ml of DEPC were added per 100ml of solution to be treated and the mixture allowed to stand overnight at 37°C before autoclaving for 20 minutes at 121°C, 15lb/in<sup>2</sup>.

**Digoxigenin - RNA labelling mix (10x)** 10mM ATP, 10mM CTP, 10mM GTP, UTP 6.5 mM and digoxigenin - UTP 3.5mM.

**Dimethyl sulphoxide** (Sigma)

**DNA size markers '1 Kb ladder'** (Gibco BRL)

**DNA size concentration markers**

$\lambda$  *Bst* E1 (Northumbria Biologicals). For DNA > 700 bp. Purchased ready cut.

**pBluescript™ II KS+ *Sau* 3A** (Stratagene). For DNA < 700 bp. Prepared by the digestion of a known concentration of pBluescript KS + with *Sau* 3A (Boehringer Mannheim) in the manufacturer's buffer followed by termination of the reaction with a one fiftieth volume of 0.5M EDTA. Stored at -20°C.

**EDTA (ethylenediamine tetraacetic acid)** (Sigma). 0.5M aqueous solution pH adjusted to 8.0 with NaOH.

**Ethanol** (Hayman Ltd). 25%, 50%, 70% and 75% (v/v) solutions made in water.

**Ethidium bromide** (Sigma). Stored as a 10 mg/ml aqueous solution at 4°C.

**Formaldehyde** (Sigma). 37% (v/v) aqueous solution pH 7.0.

**Gel denaturation solution** 1.5M NaCl, 0.5M NaOH.

**Gel neutralisation solution** 1.5M NaCl, 0.75M Tris-HCl pH7.5.

**Gelatin** (Sigma)

**Glutaraldehyde** (Sigma) Stored at -20°C. 0.2% (v/v) solutions with 4% (w/v) paraformaldehyde in PBS or PBT made fresh as required.

**Glycine** (Sigma)

**Glycogen** (Molecular biology grade, Boehringer Mannheim). Stored as 1mg/ml aqueous solution at -20°C.

**Heparin** (Sigma)

**Hybridisation buffer A (for radioactive *in situ* hybridisation)** 50% (v/v) deionised formamide, 0.3M NaCl, 20mM Tris-HCl pH 8, 5mM EDTA, 10% (w/v) dextran sulphate, 1% (v/v) Denhardt's solution, 0.5 mg/ml yeast tRNA. Stored at -20°C.

**Hybridisation buffer B (for filter hybridisation)** 5x SSC, 5x Denhardt's solution, 50% (v/v) formamide, 1% (w/v) SDS, 50mM sodium phosphate pH 6.8, 10% (w/v) dextran sulphate (m.w. 500 000), 100mg/ml sonicated salmon sperm DNA.

**Hybridisation buffer C (for rapid protocol whole mount *in situ* hybridisation)**  
50% (v/v) deionised formamide, 1.3xSSC (pH 5), 0.5% (w/v) CHAPS, 0.2% (v/v) Tween<sup>®</sup>20 , 50mg/ml yeast tRNA, 50 mg/ml heparin.

**Hydrogen Peroxide** (Sigma) Stored at 4°C.

**IPTG (Isopropyl-β-D-thiogalactopyranoside)** 100mM stock aqueous solution filter sterilised and stored at -20°C.

**Iso-amyl alcohol** (AnalaR, BDH).

**Isopropanol** (AnalaR, BDH).

**Kanamycin** (Sigma) Stock aqueous solution 35mg/ml (filter sterilised). Working concentration 50mg/ml for plates and broth.

**LB medium** (Luria - Bertani medium) 10g/l bacto-tryptone, 5g/l bacto-yeast extract, 0.25M NaCl. Sterilised by autoclaving for 20 minutes at 121°C, 15lb/in<sup>2</sup>.

**LB plates** 15g Agar per litre of LB autoclaved for 30 minutes at 121°C, 15lb/in<sup>2</sup> and allowed to cool to 50°C before pouring. Antibiotic supplements added as required.

**Ligation Buffer** 50mM Tris-HCl pH7.5, 10mM MgCl<sub>2</sub>, 20mM dithiothreitol, 1mM ATP, 50mg/ml BSA

**Lithium Chloride** 4M solution in DEPC - treated water. Filter sterilised.

**Low Gelling Temperature Agarose** (Nusieve GTG<sup>™</sup>, Flowgen).

**Magnesium sulphate** (Sigma) 10mM aqueous solution sterilised by autoclaving for 20 minutes at 121°C, 15lb/in<sup>2</sup>

**MABT** 100mM maleic acid, pH 7.5, 150mM NaCl, 1% (v/v)Tween<sup>®</sup>20 . Made fresh as required.

**Methanol** (BDH). 25%, 50%, 70% and 75% (v/v) solutions made in water.

**MOPS** 0.1M (N-morpholino) propane sulphonic acid, 40mM sodium acetate, 5mM EDTA.

**NBT** 4-Nitro blue tetrazolium chloride (Boehringer Mannheim). 100mg/ml solution in 70% (v/v) dimethylformamide. Stored away from light at -20°C.

**NTE** 0.5M NaCl, 10mM Tris-HCl, pH 8, 5mM EDTA.

**NTMT** 0.1M NaCl, 0.1M Tris-HCl pH 9.0, 0.1M MgCl<sub>2</sub>, 0.1% (v/v) Tween<sup>®</sup>20.  
Made fresh as required.

**dNTP mix** (Sodium salts, Boehringer Mannheim). 10mM stock mix of 2.5mM each of dCTP, dGTP, dATP and dTTP).

**rNTP mix** (Sodium salts, Boehringer Mannheim). 10mM stock mix of 2.5mM each of CTP, GTP, ATP and UTP).

**Paraformaldehyde** (Sigma). 4% (w/v) solution in PBS. Made fresh as required.

**PBS** 175mM NaCl, 3.4mM KCl, 0.01mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4.

**PBT** 0.1% (v/v) Triton X-100 in PBS

**PCR (polymerase chain reaction) buffer** 10mM Tris-HCl, pH 8.3, 50mM KCl and 1.6mM MgCl<sub>2</sub>.

**PEG 8000 Mix** 20% (w/v) polyethylene glycol (m.w. 8000) in a 1M NaCl solution.

**Phenol** saturated with 50mM Tris-HCl, pH 8 (Scotlab).

**Phenol/Chloroform** 1:1 mixture of phenol saturated with 50mM Tris-HCl, pH 8 and chloroform.

**Prehybridisation mix (for long protocol whole - mount *in situ* hybridisation)**  
50% (v/v) formamide, 5x SSC, 2% (w/v) Blocking Powder, 0.1% (v/v) Triton X-100, 0.5% (w/v) CHAPS, 1mg/ml yeast tRNA, 5mM EDTA, 100µg/ml sonicated herring sperm DNA, 50µg/ml heparin.

**Proteinase K** (Boehringer Mannheim). 20mg/ml stock in TE (5x) stored at 4°C or 10mg/ml stock in PBT made fresh as required.

**QuickHyb<sup>®</sup> Hybridisation solution** (Stratagene).

**Reverse Transcription Buffer** 50mM Tris-HCl pH 8.3, 40mM KCl, 1mM dithiothreitol, 6mM MgCl<sub>2</sub>, 0.1 mg/ml BSA.

**RNA size markers** 0.24-9.5 Kb ladder (Gibco BRL)

**RNA loading buffer** 40% (v/v) formamide, 1.8M formaldehyde, 1x MOPS, 1mM EDTA pH 8.0, 0.1% (w/v) bromophenol blue (Sigma), 0.1% (w/v) xylene cyanol (Sigma).

**RNase A** (DNase free, derived from bovine pancreas, Sigma) Powdered RNase A was dissolved at 10 mg/ml in 10mM Tris-HCl (pH 7.5)/ 15mM NaCl, heated to 90°C for 10 minutes and allowed to cool slowly before filter sterilisation and storage at -20°C.

**SDS** Sodium dodecyl sulphate. Stock solution 10% (w/v) in water.

**Sephadex G50** (Sigma)

**Sequencing gel loading buffer** 5:1 mixture of deionised formamide and 25mM EDTA containing 50mg/ml blue dextran (Sigma).

**Sheep serum** (Sigma). Heat treated at 60°C for 30 minutes and stored quick frozen in 1ml aliquots at -20°C.

**SM** 0.1M NaCl, 8mM MgSO<sub>4</sub>, 50mM Tris-HCl (pH 7.5), 0.1% (w/v) gelatin. Sterilised by autoclaving for 20 minutes at 121°C, 15lb/in<sup>2</sup>.

**Sodium Acetate (2M)** 2M solution in water. pH adjusted to 4.0 with acetic acid. Filter sterilised.

**Sodium Acetate (3M)** 3M solution in water. pH adjusted to 5.2 with acetic acid. Filter sterilised.

**Solution I** 50mM glucose, 50 mM Tris-HCl, pH 8, 10mM EDTA.

**Solution II** 1% (w/v) sodium dodecyl sulphate, 0.2M NaOH.

**Solution III** 3M potassium acetate pH 5.2.

**SSC (20x)** 3M NaCl, 0.3M trisodium citrate, pH adjusted to 7.0 with HCl.

**T4 Polymerase buffer** 50mM Tris-HCl pH 7.5, 10mM MgCl<sub>2</sub>, 10mM dithiothreitol, 0.5mg/ml BSA.

**TBE (1x)** 89mM Tris, 89 mM boric acid, 2.5 mM EDTA pH 8.

**TBS** 0.22M NaCl, 10mM Tris-HCl pH 8.0

**TBST** 0.05% (v/v) Tween<sup>®</sup> 20 in TBS

**TE (1x)** 10mM Tris-HCl pH 8.0, 1mM EDTA.

**TEMED** (Sigma) N, N, N', N' - Tetramethylethylenediamine. Stored at 4°C

**TE RNase** TE containing 20µg/ml RNase A (diluted from above stock).

**Terminal transferase buffer (2x)** 400mM potassium cacodylate, 50mM Tris-HCl pH 6.6, 0.5 mg/ml BSA, 3 mM cobalt chloride and 1.5mM dATP. Made as required from the 5x stock buffer provided with the enzyme (Boehringer Mannheim).

**Tetracycline** (Sigma) Stock solution 12.5 mg/ml in 70% ethanol stored at -20°C away from light. Working concentration 12.5µg/ml (plates).

**Top agarose** 7g agarose per litre of LB autocaved for 10 minutes at 121°C, 15 lb/in<sup>2</sup>. Stored at 4°C in batches and melted in a microwave oven before use.

**Transcription buffer** (for SP6, T7 and T3 bacteriophage RNA polymerases) 40mM Tris-HCl pH 8.0, 6mM MgCl<sub>2</sub>, 10mM dithiothreitol, 2mM spermidine.

**Tris-HCl [tris(hydroxymethyl)aminomethane]** Molar solutions were made in water and the pH adjusted as desired using HCl.

**Triton X-100** (Sigma)

**TSS (2x)** 20% (w/v) polyethylene glycol (m.w. 4000), 10% (v/v) dimethyl sulfoxide, 100mM MgCl<sub>2</sub> in LB. Initially this was made fresh prior to use, but subsequently it has been purchased from commercial suppliers (Epicentre Technologies).

**Tween<sup>®</sup>20** (Sigma).

**[ $\alpha$ <sup>32</sup>P]UTP** (400 Ci/mmol Amersham)

**[<sup>35</sup>S]UTP** (1000-1500Ci/mmol, Dupont).

**Wash solution I** 50% (v/v) formamide, 4x SSC, 1% (w/v) SDS

**Wash solution II** 1.25M NaCl, 0.1M Tris-HCl pH 7.5, 1% (v/v) Tween<sup>®</sup> 20

**Wash solution III** 50% (v/v) formamide, 1x SSC pH 4.5

**X-gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside)** (Stratagene). 40mg/ml solution in N,N dimethyl formamide stored at -20°C away from light .

**Yeast tRNA** (Sigma) Stored as 10mg/ml stock solution in DEPC - treated water.

### 2.1.3 Synthetic oligonucleotides

All of the synthetic oligonucleotides used except KELG and GDFX were manufactured by the NIMR Oligonucleotide Synthesis Service. They were supplied fully deprotected as ammonium hydroxide/ethanol stocks. KELG and GDFX were supplied in aqueous solution by Oswell Scientific Ltd.

#### **Plasmid vector specific**

<b>AH</b>	5' GACGTCGCATGCACGCGTACGTAAG	(25mer)
<b>KS</b>	5' CCCTCGAGGTCGACGGTATC	(20mer)
<b>M13F</b>	5' CGCCAGGGTTTTCCCAGTCACGAC	(24mer)
<b>M13R</b>	5' GGAAACAGCTATGACCATGATTAC	(24mer)

<b>pCRSF</b>	5' CCTCTAGACGTATGCTCGAGC 3'	(21mer)
<b>pCRSR</b>	5' CTTGGTACCGAGCTCGGATCC 3'	(21mer)
<b>PE</b>	5' CTGCAGGTACCGGTCCGGAATTGGG	(25mer)
<b>SK</b>	5' CGCTCTAGAACTAGTGGATC	(20mer)
<b>SP6</b>	5' TGAATTTAGGTGACACTATAGAAG	(24mer)
<b>T3</b>	5' AATTAACCCTCACTAAAGGG	(20mer)
<b>T7</b>	5' GCTCTAATACGACTCACTATAGGG	(24mer)
<b>T7a</b>	5' GTAATACGACTCACTATAGGGC	(22mer)

#### Gene specific

<b>HPRT 1A</b>	5' CCTGCTGGATTACATTAAGCACTG	(25mer)
<b>HPRT 1B</b>	5' GTCAAGGGCATATCCAACAACAAAC	(25mer)
<b>MGP 1</b>	5' CCAGTTGGTGTCAAAAGCC	(20mer)
<b>MGP 2</b>	5' CTCTCTGCTTTAAGGAGTCAG	(21mer)
<b>OC31</b>	5' TGTCCGCCCCGCATACGAGTTC	(21mer)
<b>OC32</b>	5' GGCCTCGAAGCGACAGATGG	(20mer)
<b>OC33</b>	5'CAGGCCCCCGGCTCGGGGTGCC	(22mer)
<b>IN1F</b>	5' AAGGATCCTGAGGGAACAGGTGGAGAAC	(27mer)
<b>IN2R</b>	5'AAGGATCCAAGCTGGACACGGAGTTTTTATC	(31mer)

### Family specific

<b>KELG</b>	5' CATRRTTGTTKAKYCCYARTTC	(20mer)
<b>GDFX</b>	5' AGCATRRTTKBYAAARTCWCC	(20mer)

### Differential display

<b>01</b>	5' TACAACGAGG	(10mer)
<b>02</b>	5' TGGATTGGTC	(10mer)
<b>03</b>	5' TCGGTCATAG	(10mer)
<b>04</b>	5' TACCTAAGCG	(10mer)
<b>05</b>	5' CTGCTTGATG	(10mer)
<b>06</b>	5' GATCTGACAC	(10mer)
<b>07</b>	5' GTACTGACTG	(10mer)
<b>08</b>	5' CTAGATAGCG	(10mer)
<b>dTNC</b>	5' TTTTTTTTTTTTNC	(14mer)
<b>dTNG</b>	5' TTTTTTTTTTTTNG	(14mer)
<b>dTNT</b>	5' TTTTTTTTTTTTNT	(14mer)
<b>dTNA</b>	5' TTTTTTTTTTTTNA	(14mer)

N = A, G, C or T

## Differential display band reamplification

<b>O1 Bam</b>	5' GATTGTCGGGATCCTACAACGAGG	(24mer)
<b>O2 Bam</b>	5' GATTGTCGGGATCCTGGATTGGTC	(24mer)
<b>O3 Bam</b>	5' GATTGTCGGGATCCTCGGTCATAG	(24mer)
<b>O4 Bam</b>	5' GATTGTCGGGATCCTACCTAAGCG	(24mer)
<b>O5 Bam</b>	5' GATTGTCGGGATCCCTGCTTGATG	(24mer)
<b>O6 Bam</b>	5' GATTGTCGGGATCCGATCTGACAC	(24mer)
<b>O7 Bam</b>	5' GATTGTCGGGATCCGTA CTGACTG	(24mer)
<b>O8 Bam</b>	5' GATTGTCGGGATCCCTAGATAGCG	(24mer)
<b>Eco dT</b>	5' GGCAGGGAATTCGGGTTTTTTTTT	(24mer)

## Adapters

<b>XCM1F</b>	5' CATTAATTA ACTGGCCAATTGAAATGTGGAGCT	(33mer)
<b>XCM1R</b>	5' CCACATTTCAATTGGCCAGTTAATTAATGGTA	(33mer)

## 2.2 GENERAL TECHNIQUES

### 2.2.1 Preparation and analysis of DNA and RNA

#### Extraction of nucleic Acids with phenol and chloroform

An equal volume of equilibrated phenol was added to the nucleic acid aqueous sample and the mixture vortexed in a polypropylene microcentrifuge tube prior to

centrifugation at 13000g for 15 seconds. The upper (aqueous) phase was carefully transferred to a fresh tube containing an equal volume of chloroform and this mixture was vortexed prior to centrifugation at 13000g for 15 seconds. The upper (aqueous) phase was removed and the nucleic acids precipitated.

#### **Precipitation of DNA with ethanol**

A 1/10 volume of 3M sodium acetate pH 5.2 was added to the DNA solution followed by 2.5 volumes 100% ethanol. After brief vortexing the mixture was stored in crushed dry ice for 10 minutes before centrifugation at 13000g for 5 minutes at room temperature. The DNA pellet was washed in 70% (v/v) ethanol prior to recentrifugation and removal of all of the supernatant and was resuspended in TE to the desired concentration. For nanogram quantities of DNA and for fragments <400 bp, usually 5µg yeast tRNA or 1µg glycogen was added as a carrier.

#### **Precipitation of DNA with isopropanol**

This was used when precipitating DNA from large aqueous volumes. The protocol is as above except that 1 volume isopropanol was substituted for 2.5 volumes 100% ethanol.

#### **Precipitation of DNA with polyethylene glycol (PEG)**

An equal volume of PEG 8000 mix was added to the DNA solution at room temperature. After brief vortexing the mixture was centrifuged at 13000g for 20 minutes at room temperature and the pellet washed as above prior to resuspension in TE.

#### **Precipitation of *in vitro* transcribed RNA with ethanol**

The synthesis reaction was stopped with 1µl 0.5M EDTA and 1/8 volume 4M lithium chloride was added prior to 2.5 volumes 100% ethanol. After brief vortexing the mixture was stored in crushed dry ice for 10 minutes before centrifugation at 13000g for 20 minutes at room temperature. The RNA pellet was washed in 70% ethanol prior to recentrifugation and removal of all of the supernatant and was resuspended in DEPC - treated water to the desired concentration.

#### **Quantitation of DNA and RNA in aqueous solution**

**Plasmid and genomic DNA** Spectrophotometric readings at wavelengths of 260nm, an OD<sub>260</sub> of 1 corresponding to 50µg/ml DNA.

**Gel - purified DNA fragments** A known volume of the DNA fragment was electrophoresed on an agarose minigel in tandem with a known amount of an

appropriate DNA size concentration marker. A digital photograph of the gel was taken under UV exposure and the relative fluorescence of the DNA was compared with the size markers using ImageQuant<sup>®</sup> software on a Macintosh<sup>™</sup> Power PC computer.

**RNA** Spectrophotometric readings at wavelengths of 260nm were taken, an OD<sub>260</sub> of 1 corresponding to 40µg/ml DNA.

**Synthetic oligonucleotides** Spectrophotometric readings at wavelengths of 260nm were taken, an OD<sub>260</sub> of 1 corresponding to 20µg/ml DNA.

### **Agarose gel electrophoresis of DNA**

Gels typically contained 0.8 - 1.2% (w/v) agarose in TBE with 0.5µg/ml ethidium bromide in the gel and running buffer. Electrophoresis was at 10V/cm in TBE at room temperature. For the satisfactory resolution of small (<350 bp) DNA fragments that were not radioactively labelled, the gels contained 2.0 - 3.5% (w/v) low gelling temperature (LGT) agarose and were electrophoresed at 5V/cm.

### **Agarose gel electrophoresis of RNA**

RNA gels contained 1.0 - 1.25% (w/v) agarose in MOPS and 0.7M formaldehyde. Prior to loading, the samples were heated to 65°C for 10 minutes in RNA loading buffer and the gels pre-run at 1V/cm in MOPS for 5 minutes. The gels were run in MOPS at 5V/cm for 4 - 6 hours in a fume hood with one buffer change.

### **Alkaline agarose gel electrophoresis of single stranded DNA**

Gels typically contained 1% (w/v) agarose in alkaline agarose gel buffer and were run at 0.25 - 0.5V/cm in this buffer at 4°C.

### **Extraction of DNA fragments from agarose gels**

This was only carried out for DNA that had been electrophoresed into LGT agarose. Fragments >350 bp were electrophoresed in regular agarose gels until adequate separation had occurred as assessed by UV visualisation. They were then electrophoresed into wells of 0.6% (w/v) LGT agarose that were cut into the original agarose gel. Fragments smaller than this were electrophoresed as above directly into 2.0 - 3.5% (w/v) LGT agarose gels. The areas of the gels containing the DNA fragments of interest were excised with a clean scalpel blade, trimming off as much excess agarose as possible. The LGT agarose was melted at 65°C for 15 minutes and digested for 1 hour with Gelase<sup>™</sup> (Epicentre Technologies, 1U/µg LGT agarose) at 45°C using the manufacturer's buffer followed by extraction with phenol and chloroform and ethanol precipitation of the DNA with yeast tRNA.

### **Southern blotting of plasmid DNA**

Enzymically digested plasmid DNA was electrophoresed in a 1% (w/v) agarose gel until adequate separation of DNA had occurred. A photograph was taken of the gel alongside a ruler during u.v. transillumination and it was soaked in 500ml gel denaturation solution with gentle agitation for 1 hour at room temperature. After this, it was soaked in gel neutralisation solution for 30 minutes also at room temperature and then in 20x SSC for 5 minutes. Subsequently a Southern transfer apparatus was set up as described in Sambrook *et al.* (1989) and the DNA in the gel blotted onto Hybond™-N<sup>+</sup> positively charged nylon membrane (Amersham) overnight. The next morning, the apparatus was dismantled and the membrane dried for 10 minutes at 80°C before u.v. crosslinking in a Stratalinker® u.v. crosslinker (Stratagene) at 1200µJoules.

### **Rapid alkaline Southern blotting of PCR products**

PCR reaction products were electrophoresed in a 1% (w/v) agarose gel until adequate separation of DNA had occurred. A photograph was taken of the gel alongside a ruler during u.v. transillumination and it was then soaked in 500ml 0.5M NaOH for 30 minutes. It was then blotted onto Hybond™-N<sup>+</sup> positively charged nylon membrane (Amersham) using the apparatus described by Koetsier *et al.* (1993) for 1 - 2 hours. After this, the membrane was washed in gel neutralisation solution for 5 minutes and subsequently dried for 10 minutes at 80°C.

### **Northern blotting of RNA from formaldehyde gels**

RNA was electrophoresed in a formaldehyde gel as described above. In each case 2 tracks containing 0.24-9.5 Kb RNA ladder were run, one of which contained 1µl 0.5µg/ml ethidium bromide. After the electrophoresis was complete, the migration of the ladder that had been stained with ethidium bromide was observed during u.v. transillumination and the gel was washed 3 times in 500ml DEPC - treated water for 10 minutes with gentle agitation. Subsequently, it was soaked for 45 minutes in 20x SSC. Then a Southern transfer apparatus was set up as described in Sambrook *et al.* (1989) and the RNA in the gel blotted onto Hybond™-N nylon membrane (Amersham) overnight. The next morning, the apparatus was dismantled and the membrane dried for 10 minutes at 80°C before u.v. crosslinking. The RNA ladder lanes were cut from the gel and the one containing ethidium bromide was viewed during u.v. transillumination in order to assess the efficiency of transfer. The other was stained for 15 minutes in 20ml 0.5M sodium acetate pH 5.2, 0.04% (w/v) methylene blue and destained in 20ml DEPC - treated water for 15 minutes. The size

markers that had transferred were now clearly visible on the filter and were marked for future reference

### **2.2.2 *E. coli*, plasmids and bacteriophages**

#### **Preparation of competent bacterial host cells**

Using a sterile wire the *E. coli* strain required was streaked from a frozen stock onto an LB - agar plate and incubated at 37°C overnight. A single colony of cells was picked and grown in a 100ml LB culture to an OD<sub>600</sub> of 0.4 (corresponds to 10<sup>8</sup> cells/ml). 100ml ice cold TSS(2x) was added to the culture and the two mixed gently on ice. 200µl aliquots were frozen in a dry ice/ethanol bath and stored long term at -70°C. Typical transformation frequencies obtained using this method were of the order of 10<sup>7</sup> - 10<sup>8</sup> colonies/µg DNA.

#### **Chemical transformation of competent bacterial host cells**

Competent cells were thawed in ice for 15 minutes and approximately 10 - 100ng plasmid DNA were added either in the form of aliquots of ligation reactions or plasmid DNA preparations. After incubation on ice for a further 15 minutes the cells were heat - shocked at 42°C for 60 seconds. 1ml BHI was added and the cells incubated at 37°C for 30 minutes with agitation at 225 rpm. Several volumes of this culture were plated on appropriate antibiotic containing plates and incubated overnight at 37°C. *Lac* screening of plasmids was achieved by the addition of 40µl X-gal and 4µl IPTG stock solutions per culture plate.

#### **Electro-transformation of commercially acquired competent cells**

Ligation reactions were brought to 50µl total volume with water. After ethanol precipitation with 1µg glycogen, the pellets were washed twice in 70% (v/v) ethanol prior to resuspension in 10µl deionised water. 1µl each ligation was mixed with 10µl electrocompetent XL-1 Blue cells (purchased directly from Stratagene) in pre-chilled 0.1cm electroporation chambers (Biorad). Electroporations were then performed using a Gene Pulser (Biorad) set at 1.7kV, 200Ω, 25µF with a pulse length of 4.7msec. After each electroporation, 1ml pre-warmed (37°C) BHI was added and the cells incubated for 30 minutes at 37°C. 10µl were plated onto LB-ampicillin plates and incubated overnight.

#### **Preparation of bacteria for plating of λ bacteriophage**

Using a sterile wire XL1-Blue cells were streaked from a frozen stock onto an LB - agar plate and incubated at 37°C overnight. A single colony of cells was picked and grown overnight in a 50ml LB culture supplemented with 0.2% (w/v) maltose. The

culture was centrifuged at 4000g for 10 minutes and the supernatant discarded. The cell pellet was resuspended in 10mM magnesium sulphate to OD<sub>600</sub> of 2.0. The cell suspension was stored at 4°C for up to two weeks.

### **Plating of $\lambda$ bacteriophage**

Original bacteriophage stocks were diluted tenfold serially in SM. 0.1ml each dilution were mixed with 0.1ml plating bacteria and incubated at 37°C for 20 minutes. 3ml molten (47°C) top agarose were added and the mixture quickly poured onto a small culture plate which was quickly swirled in order to evenly distribute the top agarose. The plates were then incubated at 37°C overnight and the number of plaques that appeared was counted in order to estimate the number of clones per ml in each of the library dilutions.

### **Large scale plasmid DNA preparations from bacterial cultures**

200ml bacterial cultures containing the appropriate antibiotic were spun down in a Sorvall™ RC5C refrigerated centrifuge with a GSA rotor for 10 minutes at 4000rpm at 4°C. The pelleted cells were resuspended in 20ml solution I and 40 ml solution II were added with gentle agitation. The mixture was left at room temperature for 5 minutes before the addition of 30ml solution III. The mixture was immediately agitated before being spun once more at 4°C in the GSA rotor for 5 minutes at 8000rpm. The supernatant was filtered through gauze dressing into 40 ml isopropanol and the new mixture was agitated before a final spin at 4°C in the GSA rotor for 10 minutes at 10000rpm. The resulting pellet was carefully dried before resuspension in 5ml TE. 4.75ml of the dissolved pellet was added to a Sterilin™ tube containing 5.1gm caesium chloride. After the caesium chloride had dissolved, 0.25ml ethidium bromide (10 mg/ml) were added and any proteinaceous precipitate was spun down. The supernatant was transferred into an appropriately sized QuickSeal™ tube (Beckman) before spinning at room temperature for 4 hours using a VTI-80 rotor in a Beckman L8-M ultracentrifuge at 70000 rpm. The resulting high quality plasmid DNA was removed from the centrifuge tube as described (Sambrook *et al.*, 1989). Ethidium Bromide was extracted from the DNA sample by mixing with 4 volumes of water saturated butanol. Using a pasteur pipette, the lower aqueous phase was removed and transferred to a fresh container and the extraction step was repeated. The DNA was then extracted with phenol followed by chloroform and precipitated with 100% ethanol at room temperature. The resulting pellet was washed with 70% (v/v) ethanol and redissolved in TE at 1µg/ml.

### **Small scale plasmid DNA preparations from bacterial cultures**

1ml bacterial cultures containing the appropriate antibiotic were spun down in microcentrifuge tubes at 10000g for 4 minutes and the resulting bacterial pellets resuspended in 100µl solution I. 250µl solution II were then added with brief agitation followed by 250µl solution III. This was followed by 500µl phenol/chloroform and vortexing before centrifugation at 13000g for 4 minutes at room temperature. 400µl of the resulting upper aqueous phases were transferred to fresh tubes for isopropanol precipitation. The resulting washed pellets (enriched in plasmid DNA) were resuspended in 50µl TE RNase.

### **2.2.3 General radiolabelling techniques**

#### **Preparation of radioactive size markers**

250ng pBluescript<sup>TM</sup> II KS<sup>+</sup> vector (Stratagene) were digested with 20U *Sau3A* (Boehringer) for 60 minutes at 37°C in a 20µl reaction volume. 50µCi (1.5mM) [ $\alpha^{32}\text{P}$ ]dATP was added directly to this along with 1/10 volume of a 5mM each dGTP, dCTP, dTTP and 1U Klenow fragment (New England Biolabs). The reaction was incubated at 30°C for 15 minutes and then stopped with 2.5µl 0.5M EDTA. Typically a 5µl aliquot was loaded onto differential display gels.

#### **Preparation of $\alpha^{32}\text{P}$ -labelled riboprobes for filter hybridisation**

The protocol was based upon that of Melton *et al.* (1984). 20µg plasmid bearing the clone of interest was digested with an appropriate restriction enzyme and subsequently purified by phenol extraction and ethanol precipitation followed by resuspension in TE at 1µg/µl. Probes were synthesised using 1µg template, 70µCi [ $\alpha^{32}\text{P}$ ]UTP, 10U of the appropriate RNA polymerase (SP6, T3 or T7, Boehringer), 0.25 mM each GTP, CTP and ATP, 1.875µM unlabelled UTP and transcription buffer in 20µl reaction volumes. Routinely, the reactions contained 25U Rnasein<sup>TM</sup> (human placental Rnase inhibitor, Promega) and they were incubated for 1 hour at 37°C. Subsequently the reactions were treated with 1U RQ 1<sup>TM</sup> RNase free DNase (Promega) for 10 minutes before being made up to 100µl volumes with DEPC - treated water. The riboprobes were extracted once with phenol and then with chloroform prior to precipitation with 4M lithium chloride as described above. Washed riboprobe pellets were redissolved in 100µl DEPC - treated water and 2 volumes of 100% ethanol were added. The probes were stored in this state at -70°C until required, at which time 0.1 volume 4M lithium chloride was added and they were precipitated.

### **Preparation of <sup>35</sup>S-labelled riboprobes for *in situ* hybridisation**

This was as described by Wilkinson and Green, (1990). 20µg plasmid bearing the clone of interest was digested with an appropriate restriction enzyme and subsequently purified by phenol extraction and ethanol precipitation followed by resuspension in TE at 1µg/µl. Probes were synthesised using 1µg template, 50µCi [<sup>35</sup>S]UTP, 10U of the appropriate RNA polymerase (SP6, T3 or T7, Boehringer), 2.5 mM each GTP, CTP and ATP and transcription buffer in 20µl reaction volumes. Routinely, the reactions contained 25U Rnasein<sup>TM</sup> (human placental Rnase inhibitor, Promega) and they were incubated for 1 hour at 37°C. Subsequently they were treated with 1U RQ 1<sup>TM</sup> RNase free DNase (Promega) for 10 minutes before fractionation on a Sephadex G50 column set up in a 1ml sterile disposable syringe with a 300µl elution volume in column wash buffer. The column eluates were precipitated with ethanol and the precipitated riboprobe/tRNA pellets were washed twice with 80% (v/v) ethanol and once with 100% ethanol before resuspension in 10µl 100mM DTT. 1µl of this was counted in a scintillation counter (Beckman) and the solution was dissolved to 2x10<sup>5</sup> dpm/µl in a 9:1 (by volume) mixture of 100mM DTT and hybridisation buffer. Newly made riboprobes could be stored indefinitely at -70°C in the hybridisation buffer A.

### **Random primer - labelling of DNA probes for filter hybridisation**

Gel purified DNA templates were generated by PCR or by restriction digestion. 25ng of purified template were then used for random primer - labelling with the Prime-it<sup>®</sup>II kit (Stratagene) using the manufacturer's protocol. 50µCi [ $\alpha$  <sup>32</sup>P]dCTP were used per reaction. The reaction products were extracted once with phenol and once with chloroform prior to ethanol precipitation with yeast tRNA. The precipitated probe was washed and resuspended in 20µl water. 1µl of this was used for scintillation counting and specific activity estimates were performed as described below. Probes were used at specific activity greater than 10<sup>8</sup> dpm/µg.

### **Estimation of the specific activity of radioactively labelled probes**

Each method assumes that 1µCi = 2.2x10<sup>6</sup> dpm (Sambrook *et al.*, 1989) and that no free radioactivity is left with the probe after whatever purification procedures are adopted. The activity of 1µl of purified probe was measured directly in a liquid scintillation counter (Beckman).

$$\text{specific activity (dpm/}\mu\text{g)} = \frac{\text{total activity incorporated (dpm)}}{\text{probe yield } (\mu\text{g})}$$

### **DNA probes synthesised by random primer - labelling**

During this process there is net DNA synthesis while the initial substrate remains unlabelled. Both participate in the hybridisation.

$$\text{Probe yield} = \text{ng initial substrate DNA} + \text{ng DNA synthesised}$$

Given that the mean molecular weight of a nucleoside monophosphate in DNA is 340 and that there are 4 nucleotides incorporated into the probe, only one of which is labelled, then for a labelled nucleotide of activity  $\kappa \times 10^3 \text{ Ci/mmol}$

$$\text{ng DNA synthesised} = \frac{\mu\text{Ci incorporated} \times 0.35 \times 4}{\kappa}$$

### **Probes synthesised by bacterial RNA polymerases**

During this process there is net RNA synthesis while the initial substrate is destroyed and does not participate in the hybridisation reaction. Consequently, probe yield depends on the percentage of label that is incorporated and specific activity only on the specific activity of the labelled nucleotide in the synthesis reaction.

$$\text{Probe yield} = \% \text{ incorporation} \times \text{maximum theoretical RNA yield (ng)}$$

$$\% \text{ incorporation} = \frac{\text{incorporated dpm}}{\text{total dpm}}$$

$$\text{maximum theoretical RNA yield (ng)} = \text{total nmol limiting rNTP} \times 4 \times 340$$

(Given that the mean molecular weight of a nucleoside monophosphate in DNA is 340 and that there are 4 nucleotides incorporated into the probe, only one of which is labelled).

#### **2.2.4 Filter hybridisation**

All Southern and Northern blot hybridisations were carried out in a Hybaid oven using roller bottles. Colony lift filters were hybridised in 22cm x 22cm plastic culture trays and were washed in an agitating water bath (Kotterman).

### **Southern blots and colony lifts (*vide infra*)**

DNA probes were used. Hybridisations were carried out in QuickHyb<sup>®</sup> hybridisation solution (Stratagene) using the manufacturer's protocol. For probes < 300bp, hybridisation was carried out at 60°C with 3 washes at room temperature and 1 at 60°C all for 5 minutes in 0.5x SSC, 0.5% (w/v) SDS. For longer probes, hybridisation was carried out at 65°C followed by 3 washes at room temperature and 1 at 65°C all for 5 minutes in 0.1x SSC, 0.5% (w/v) SDS. After hybridisation, Southern blot filters were wrapped in Saran Wrap<sup>™</sup> before being exposed to a Phosphor screen (Molecular Dynamics Inc.). The screens were analysed on a Phosphor-imager using ImageQuant<sup>®</sup> software on a Macintosh<sup>™</sup> Power PC computer.

### **Northern blots**

RNA or DNA probes were used. Filters were prehybridised for 2 hours in 0.5ml hybridisation buffer B per cm<sup>2</sup> of filter at 42°C. Hybridisation solution was then prepared by either boiling a DNA probe or heating an RNA probe to 80°C for 5 minutes before mixing with the same volume hybridisation buffer B as was used for prehybridisation. (For probes of specific activity 10<sup>8</sup> dpm/μl and greater, a probe concentration of approximately 5ng/ml was used). The prehybridisation mixture was decanted and immediately replaced with the hybridisation solution and the filter hybridised overnight at 42°C for DNA probes or 65°C for RNA probes. After hybridisation the filters were washed 3 times at room temperature in 0.5x SSC, 0.5% (w/v) SDS for 5 minutes followed by 1 wash in the same solution at 68°C for 15 minutes. After hybridisation, filters were wrapped in Saran Wrap<sup>™</sup> before being exposed to a Phosphor screen (Molecular Dynamics Inc.). The screens were analysed on a Phosphor-imager using ImageQuant<sup>®</sup> software on a Macintosh<sup>™</sup> Power PC computer.

## **2.3 SPECIFIC PROTOCOLS**

### **2.3.1 Collection of mouse embryos**

Embryos were obtained from naturally mated CBA/N mice in the NIMR animal facility. All dissections were performed in ice-cold PBS.

### **2.3.2 RNA isolation and purification from mouse tissues**

Total RNA was prepared by the acid-guanidinium-phenol method (Chomczynski and Saatchi, 1987). 1ml denaturing solution was added per 100mg of tissue at room temperature and the mixture homogenised. Then a 1/10 volume 2M sodium acetate and an equal volume water saturated (unbuffered) phenol were added with vortexing.

The mixture was left at room temperature for 5 minutes before the addition of 0.2ml chloroform/iso-amyl alcohol per ml denaturing solution with further vortexing. After 15 minutes on ice, it was spun at 13200g for 15 minutes at 4°C. The aqueous phase (containing RNA) was removed with care to avoid the interface (containing DNA and protein). Total RNA was precipitated with isopropanol, resuspended in DEPC - treated water to the required concentration and stored at -70°C.

### **2.3.3 Purification of poly (A)<sup>+</sup> RNA from total RNA**

Oligotex-dT<sup>TM</sup> latex beads (Qiagen) were used as directed by the manufacturer. All poly (A) <sup>+</sup> enriched RNA was stored at -70°C.

### **2.3.4 In vitro synthesis of representative RNA from directionally cloned cDNA plasmid libraries**

10µg of plasmid library (purified by ultracentrifugation) was linearised by digestion with *NotI* (Boehringer Mannheim), extracted once with phenol and then with chloroform prior to ethanol precipitation and resuspension in DEPC - treated water at 1µg/µl. 1µl of this was used in a 40µl transcription reaction with transcription buffer, 500µM each rNTP, 25U Rnasein<sup>TM</sup> (human placental Rnase inhibitor, Promega) and 10U T7 RNA polymerase (Boehringer Mannheim). The reaction was allowed to proceed for 2 hours at 37°C after which 5U RQ 1<sup>TM</sup> RNase free DNase (Promega) were added and the reaction incubated for a further 45 minutes at 37°C. The newly synthesised RNA was precipitated with LiCl and resuspended in DEPC - treated water at 1µg/µl as assessed by spectrophotometry. A 1µl aliquot was electrophoresed in an agarose/TBE gel in order to ascertain an approximate guide to the integrity of the synthesised RNA.

### **2.3.5 Synthesis of cDNA from dissected somites and presomitic mesoderm**

Poly (A) <sup>+</sup> enriched RNA was purified from freshly dissected first-defined somites and presomitic mesoderm from 8.0 *dpc* and 9.0 *dpc* embryos as described above. The RNA derived from the equivalent of 5 of each was mixed with 100ng dT<sub>24</sub> in a volume of 11µl DEPC - treated water and heated to 65°C for 10 minutes. After incubation at room temperature for a further 15 minutes, the mixture was made to 19µl with reverse transcription buffer and 500µM each dNTP. Following a further incubation for 5 minutes at 37°C, 100U (0.5µl ) Superscript<sup>TM</sup> (Moloney murine leukaemia virus reverse transcriptase, GIBCO-BRL) were added and the reaction incubated for 1 hour at 37°C. The reaction mixture containing newly synthesised cDNA was then placed at 65°C for 10 minutes. A parallel reaction was performed using 1/5th the amount of starting material which was labelled by the addition of

2.5 $\mu$ Ci (75 nM) [ $\alpha^{33}$ P]dATP, these radioactive reaction products were electrophoresed in an alkaline agarose gel and Southern blotted for autoradiography.

### **2.3.6 Reverse transcription - PCR analysis of embryonic RNA**

Poly (A) + enriched RNA derived from the equivalent of one 8.5 *dpc* or three 7.5 *dpc* embryos was mixed with 250ng dT<sub>24</sub> in a volume of 11 $\mu$ l DEPC - treated water and heated to 65°C for 10 minutes. After incubation at room temperature for a further 15 minutes, the mixture was made to 19 $\mu$ l with reverse transcription buffer and 200 $\mu$ M each dNTP. Following a further incubation for 5 minutes at 37°C, 200U (1 $\mu$ l) Superscript<sup>TM</sup> were added and the reaction run for 1 hour at 37°C. The reaction mixture containing newly synthesised cDNA was then placed at 65°C for 10 minutes and either stored at -20°C or used directly in PCR. PCR reactions were performed directly on 2 $\mu$ l aliquots of the reverse transcription reaction mixtures in 20 $\mu$ l volumes with PCR buffer, 200 $\mu$ M each dNTP, 0.5 $\mu$ M each primer and 2U AmpliTaq<sup>®</sup> DNA polymerase (Perkin-Elmer). When using gene specific primers, 30 cycles of 95°C for 30 seconds, 55°C for 20 seconds and 73°C for 30 seconds were performed followed by 5 minutes at 73°C. Reaction products were visualised under u.v. exposure after agarose gel electrophoresis.

### **2.3.7 mRNA differential display (conventional protocol)**

#### **Templates**

**Adult tissue RNA** Reverse transcription was performed on 1 $\mu$ g poly (A)<sup>+</sup> enriched RNA from each adult tissue.

**Embryonic RNA** Reverse transcription was performed on poly (A)<sup>+</sup> enriched RNA obtained from an equivalent of fifteen 7.5 *dpc* embryos or two 8.5 *dpc* embryos.

**In vitro synthesised RNA** Reverse transcription was performed on 1 $\mu$ g synthetic RNA from each library.

#### **Reverse Transcription**

Each RNA was mixed with 2.5 $\mu$ M dTNC, dTNG, dTNT or dTNA as appropriate in 11 $\mu$ l DEPC - treated water and heated to 65°C for 10 minutes. After incubation at room temperature for a further 15 minutes, the mixture was made to 19 $\mu$ l with reverse transcription buffer and 25 $\mu$ M each dNTP. Following a further incubation for 5 minutes at 37°C, 300U (1.5 $\mu$ l) Superscript<sup>TM</sup> (GIBCO BRL) were added and the reaction run for 1 hour at 37°C. The reaction mixture containing newly synthesised cDNA was then placed at 65°C for 10 minutes.

## PCR

Each cDNA synthesis reaction was then divided into 8 aliquots of 2.5µl and used in 20µl PCR reactions in PCR buffer with 2.5µM reverse transcription primer, 0.5µM one of O1 to O8, 2.5µM each dNTP, 2.5U AmpliTaq® DNA polymerase (Perkin-Elmer) and 2.5µCi (75 nM) [ $\alpha^{33}\text{P}$ ]dATP. In the case of somite and presomitic mesoderm cDNA, 2.5µl gel purified cDNA was used. The reactions were performed with the parameters 94°C for 90 seconds, followed by 40 cycles of 94°C for 30 seconds, 40°C for 60 seconds, 73°C for 45 seconds and an additional extension period of 5 minutes at 73°C.

### 2.3.8 mRNA differential display (modified protocol)

#### Templates

**Somite and presomitic mesoderm cDNA** Reverse transcription was initially performed on poly (A)<sup>+</sup> enriched RNA obtained from an equivalent of one dissected presomitic mesoderm tissue block or five dissected somites. Subsequent experiments used a quarter of the poly (A)<sup>+</sup> enriched RNA obtained from individually isolated structures dissected from the same embryo.

#### Reverse Transcription

Each RNA was mixed with 1µM dT<sub>24</sub> in 11µl DEPC - treated water and heated to 65°C for 10 minutes. After incubation at room temperature for a further 15 minutes, the mixture was made to 19µl with reverse transcription buffer and 25µM each dNTP. Following a further incubation for 5 minutes at 37°C, 300U (1.5µl ) Superscript™ (GIBCO BRL) were added and the reaction run for 1 hour at 37°C. The reaction mixture containing newly synthesised cDNA was then placed at 65°C for 10 minutes.

## PCR

Each cDNA synthesis reaction was then divided into 8 aliquots of 2.5µl and used in 20µl PCR reactions in PCR buffer containing 0.5µM one of the eight reamplification 24-mer primers, 2.5µM each dNTP, 2.5U AmpliTaq® DNA polymerase (Perkin-Elmer) and 2.5µCi (75 nM) [ $\alpha^{33}\text{P}$ ]dATP. The reactions were performed with the parameters 94°C for 90 seconds, followed by 2 cycles of 94°C for 30 seconds, 40°C for 180 seconds, 73°C for 90 seconds and subsequently 40 cycles of 94°C for 30 seconds, 60°C for 60 seconds, 73°C for 90 seconds and an additional extension period of 5 minutes at 73°C.

### **2.3.9 Gel electrophoresis of differential display products**

In all cases, reaction mixtures from the different tissues under investigation which had been generated with the same primer pairs were electrophoresed in adjacent tracks of the gel. All gels (both native and denaturing) were 0.4mm thick and poured in 30cm plates.

**Denaturing gel electrophoresis** 5µl of each PCR reaction were mixed with 1µl RNA loading buffer and the mixture heated to 90°C for 5 minutes prior to loading onto a 6% (w/v) denaturing (7M urea) polyacrylamide gel in TBE buffer. Electrophoresis was performed for 2 hours at 40 V/cm, by which time the bromophenol blue dye would have reached the bottom of the gel. No size markers were used but the approximate sizes of reaction products were inferred from the migration of the dyes in the loading buffer as previously described (Sambrook *et al.*, 1989). Subsequently, the gels were fixed in 5% (v/v) methanol 5% (v/v) acetic acid and dried for autoradiography using Kodak™ BIOMAX-MR X-ray film at room temperature for 24 hours.

**Native gel electrophoresis** 5µl of each PCR reaction were mixed with 1µl agarose gel loading buffer and the mixture subsequently loaded onto a 6% (w/v) polyacrylamide gel and electrophoresed in TBE buffer at 4°C for 16 hours at 300V. Approximately 5ng pBluescript KS + *Sau3A* radiolabelled size marker were loaded to each gel. Subsequently the gels were dried for autoradiography using Kodak™ BIOMAX-MR X-ray film at room temperature for 24 hours.

### **2.3.10 Elution and reamplification of DNA bands from native differential display gels**

#### **Conventional Protocol**

The dried gels were triangulated with respect to the X-ray film with Glogos™ II autoradiographic markers (Stratagene), and the two were carefully taped together. Products that appeared to be of interest were then excised using a scalpel blade. The filter paper was gently teased off the back of the dried polyacrylamide and the latter was left to soak in 100µl water at room temperature for 4 hours. A 5µl aliquot of this was used for subsequent reamplification PCR. In each case, the PCR primers used were Eco dT and the 5' extended reamplification primer corresponding to the random primer that was used in the original differential display reaction. For example, if O1

was used in the differential display, then Bam O1 was used for reamplification. The reamplification reactions were performed in 40µl volumes with each primer at 1µM and dNTPs at 25µM. The cycling parameters were 2 minutes at 94°C followed by 2 cycles of 94°C for 30 seconds, 40°C for 60 seconds, 73°C for 60 seconds and then 34 cycles of 94°C for 20 seconds, 58°C for 30 seconds, 73°C for 30 seconds followed by a final extension for 5 minutes at 73°C. All other parameters were as for the differential display except that no radiolabel was added.

### **Modified protocol**

The dried gels were triangulated with respect to the X-ray film with Glogos<sup>TM</sup> II autoradiographic markers (Stratagene), and the two were carefully taped together. Products that appeared to be of interest were then excised using a scalpel blade. The filter paper was gently teased off the back of the dried polyacrylamide and the latter was left to soak in 100µl water at room temperature for 4 hours. A 5µl aliquot of this was used for subsequent reamplification PCR. These reactions were carried out in 40µl volumes with the same primer and dNTP concentration as the differential display but in the absence of radiolabel. The cycling parameters were 60 seconds at 94°C followed by 40 cycles of 94°C for 30 seconds, 60°C for 30 seconds, 73°C for 90 seconds followed by a final extension for 5 minutes at 73°C

### **2.3.11 Cloning of reamplified differential display fragments**

#### **T/A Cloning vectors**

**pCR<sup>TM</sup> II** (Invitrogen). This was purchased from the manufacturer and used according to the manufacturer's instructions.

**pTAx**. This was derived from pBluescript<sup>TM</sup> II KS<sup>+</sup> (Stratagene) and made in the laboratory. 500ng pBluescript<sup>TM</sup> II KS<sup>+</sup> were digested with *Kpn*I and then *Sac*I (Boehringer Mannheim) and the vector backbone purified at 50 ng/µl in TE after agarose gel electrophoresis. 2.5µg each adapter oligonucleotide XCM 1F and XCM 1R were heated to 90°C in 10µl TE and the mixture allowed to cool to room temperature over 4 hours. 1µl vector backbone solution and 1µl annealed adapter mixture were then added to a 10µl ligation reaction containing 5U T4 DNA ligase (Boehringer Mannheim) in ligation buffer. After 2 hours incubation at room temperature, 1µl of the reaction was used to transform Epicurian Coli<sup>TM</sup> XL1-Blue bacterial cells (Stratagene) and colonies were selected on the basis of resistance to ampicillin. Small scale plasmid DNA preparations were made from 12 such clones and those which could be digested with *Msc*I (New England Biolabs) were chosen as being

recombinant for the adapter linker. A large scale plasmid DNA preparation was made from one of these and 10µg of this was digested with *Xcm* I (New England Biolabs) in a 250µl reaction using the manufacturer's buffer. The resulting T/A cloning vector was purified by phenol and chloroform extraction and stored in TE at 50ng/µl.

#### **Blunt - ended cloning vector**

**pBluescript<sup>TM</sup> - *EcoRV***. This was prepared in the lab by the digestion of pBluescript<sup>TM</sup> II KS<sup>+</sup> (Stratagene) with *EcoRV* (Boehringer Mannheim) in the manufacturer's buffer. It was gel purified and stored in TE at 50ng/µl.

#### **T/A ligation**

Whatever the vector, T/A cloning was achieved by adding 7µl of the unpurified reamplification mixture directly into a 10µl ligation reaction containing 50ng vector and 5U T4 DNA ligase (Boehringer Mannheim). Reactions were incubated for 4 hours at 16°C.

#### **Blunt - ended ligation**

At the end of the initial reamplification reaction, 2.5U cloned *Pfu* polymerase (Stratagene) and 4µl of 2mM dNTPs were added directly to the *Taq* reaction below the mineral oil and the mixture was incubated at 73°C for 45 minutes. 7µl of this "*Pfu* polishing" reaction were added directly to a ligation reaction outlined above but containing 50ng pBluescript - *EcoRV*.

#### **Bacterial transformation and PCR screening of recombinants**

Ligation reactions were transformed into NM522 cells. Transformed bacteria were plated onto ampicillin plates and *lac* screening was used to identify potentially recombinant clones (white colonies). White or pale blue colonies were picked from the plates and grown in 100µl volumes of sterile BHI (containing 25 mg/ml ampicillin) gridded on 96 well tissue culture trays (Costar) for 2 hours at 37°C. A PCR mix containing 0.25µM each appropriate reamplification primer, 50µM each dNTP and 1U AmpliTaq<sup>®</sup> DNA polymerase (Perkin-Elmer) per 20µl reaction was dispensed into a corresponding 96-V Thermowell<sup>TM</sup> thin - walled PCR tray (Costar) and the gridded cultures were picked and touched from the tissue culture tray into it. This was achieved using a sterilised "hedgehog" device made from a perspex block with 96 correctly spaced protruding steel spikes. The PCR tray was overlain with a heat sealable film (MJ Research) and cycled with the following parameters; 92°C for 2

minutes, 25 cycles of 94°C for 20 seconds, 58°C for 20 seconds and 73°C for 30 seconds followed by a 5 minute extension period of 73°C. The reaction products were visualised after agarose gel electrophoresis.

### **2.3.12 Library screening**

#### **Plasmid libraries**

The transformation efficiency of the library was assessed in XL - 1 Blue cells. A transformation was then carried out (in 1ml total volume of BHI medium) with the aim of giving 8 - 10 times as many clones as were thought to be in the initial library. The recovery phase of the transformation was allowed to proceed for 15 minutes only before it was split into 48 20µl aliquots that were grown overnight as 1ml miniculture pools in BHI medium (containing 25 mg/ml ampicillin). 1µl, 5µl and 10µl aliquots of the remainder of the transformation were plated onto ampicillin plates in order to further verify the number of clones per pool. Small scale plasmid preparations were made from each pool. These were digested, electrophoresed in agarose gels and subsequently Southern blotted. The filters were hybridised with radiolabelled probes and pools which gave a positive signal were selected for transformation as above but this time the transformation was performed with the aim of giving 2 - 3 times as many clones as were thought to be in the initial library pool. When a clone was thought to be in a pool of approximately 100 clones, that pool was transformed into NM 522 cells and the transformation plated. Individual colonies were picked and grown in 100µl volumes of sterile BHI (containing 25 mg/ml ampicillin) gridded on 96 well tissue culture trays as described above for 2 hours at 37°C. A PCR mix containing 0.25µM each M13F and M13R oligonucleotide primers, 50µM each dNTP and 1U AmpliTaq<sup>®</sup> DNA polymerase (Perkin-Elmer) per 20µl reaction was dispensed into a corresponding 96-V Thermowell<sup>™</sup> thin - walled PCR tray (Costar) and the gridded cultures were picked and touched from the tissue culture tray into it using the hedgehog device. The PCR tray was overlain with a heat sealable film (MJ Research) and cycled with the following parameters; 92°C for 2 minutes, 30 cycles of 94°C for 40 seconds, 58°C for 20 seconds and 73°C for 2 minutes followed by a 15 minute extension period of 73°C. The reaction products were electrophoresed in an agarose gel and Southern blotted. The filter was hybridised with the probe of interest and positive clones were identified for further study.

#### **Bacteriophage libraries**

Libraries were titred as described above (see "Plating of λ bacteriophage") after serial dilution in SM and infection of freshly prepared XL 1 Blue plating cells. 10<sup>5</sup> phage and 1ml plating cells were mixed and incubated at 37°C for 20 minutes in a sterile 50

ml polypropylene bottle. 20 ml top agarose (47°C) were added and after gentle swirling the mixture was poured onto a 22cm x 22cm LB plate. Library plates were allowed to stand for 10 minutes at 37°C before being inverted and incubated (at 37°C) for 12 - 14 hours. After this they were stored at 4°C for 2 hours before colony lifts were attempted. 22cm x 22cm sheets of Hybond™ - N<sup>+</sup> filter were carefully overlain onto the library plates and left for 1 minute at room temperature. Each filter and the agar were triangulated with a sterile needle and the plastic casing of the plate was marked in the corresponding positions with a permanent marker. Then it was removed from the library plate and placed colony side up on a pad of filter paper soaked in gel denaturing solution and left for 7 minutes. Subsequently it was neutralised for 3 minutes in a similar manner with gel neutralisation solution before a final wash in 2 x SSC. After this, it was fixed as described above using a UV crosslinker. All library plates were lifted in duplicate. Library filters were hybridised with the probe of interest as described above. Positive clones as defined after autoradiography were confirmed if the duplicate filters showed corresponding positives and were then cored from the library plate with the end of a 1ml pipette tip (Rainin). The cored positive pick was resuspended in 0.5ml SM containing 40µl chloroform and eluted from the top agarose by incubation at room temperature for 4 hours. Tenfold serial dilutions were then made in SM and these were used to infect fresh plating cells as described above, except that colonies were plated onto small LB dishes. Plates from dilutions which gave clearly separated colonies were chosen for duplicate lifts and the resulting filters were hybridised as described. Clearly defined positive plaques were cored and the phage eluted as described. In the case of Lambda ZAP® vectors, corresponding pBluescript™ SK plasmid was excised as described below.

#### **Excision of Bluescript™ from Lambda ZAP® vectors**

XL -1 Blue MRF' plating cells were prepared as described. In addition an overnight culture of a single colony of SOLR cells was grown at 30°C in 50ml LB without supplements. The culture was gently spun at 1000g for 5 minutes and resuspended to OD<sub>600</sub> in 10mM MgSO<sub>4</sub>. 200µl XL-1 Blue MRF' plating cells and 250µl single pick phage stock were mixed and incubated with 1µl ExAssist helper phage (>1 x 10<sup>6</sup> pfu/µl) (Stratagene) at 37°C for 15 minutes. 3ml LB were added and the mixture incubated in a sterile 50ml polypropylene tube for 2 hours at 37°C with agitation. After this it was heated to 70°C for 15 minutes before centrifugation at 4000g for 15 minutes. The supernatant was removed and 50µl were used to infect 200µl SOLR cells at 37°C for 15 minutes. 200µl of this infection were plated onto LB - ampicillin

culture dishes and incubated overnight at 37°C. 5 - 8 colonies were picked and screened by PCR as described above using M13F and M13R primers.

### **2.3.13 "Shotgun" cloning of full - length cDNA clones prior to sequencing**

50µg plasmid DNA which had been purified by ultracentrifugation was digested with appropriate restriction enzymes to release the cloned cDNA insert, and this was gel purified as described above. It was resuspended in TE at approximately 1µg/ml and 4 - 5µg were self ligated in a 10µl reaction with ligase buffer and 5U T4 DNA ligase (Boehringer Mannheim) for 1 hour at room temperature. The ligation reaction was diluted to 50µl in the microcentrifuge tube and 2µl were visualised by U.V. transillumination after agarose gel electrophoresis. The rest was sonicated by placing the microcentrifuge in a polystyrene support floating in 1cm water in a 7" cup horn vessel using a Vibra Cell VC600 High Intensity Sonicator (Sonics @ Materials Inc.). Sonication was for 60 seconds at 40 cycles. The sonicated DNA solution was made up to 100µl with the appropriate enzyme buffer and the DNA digested with 10U *EcoRV* (Boehringer Mannheim) for 1 hour at 37°C. 2µl were then visualised by U.V. transillumination after agarose gel electrophoresis in order to ensure adequate fragmentation had occurred. The remainder was precipitated with PEG 8000 and the pellet resuspended after washing in 20µl T4 polymerase buffer, 200µM each dNTP and 2U T4 DNA polymerase (New England Biolabs.). The reaction was incubated at 12°C for 20 minutes and then at 80°C for 10 minutes before repeat PEG 8000 precipitation. The washed pellet was resuspended in a 10µl reaction volume containing ligase buffer, 5U T4 DNA ligase (Boehringer Mannheim) and 50ng pBluescript<sup>TM</sup> - *EcoRV*. The ligation reaction was allowed to proceed for 4 hours at 16°C and was then made up to 20µl in the appropriate buffer and cut back with 10U *EcoRV* (Boehringer Mannheim) for 30 minutes at 37°C. 2.5µl cut back ligation reaction were then transformed into NM522 cells and the transformation plated onto LB - ampicillin plates with *lac* screening. White or pale blue colonies were picked from the plates and grown in 100µl volumes of sterile BHI (containing 25 mg/ml ampicillin) gridded on 96 well tissue culture trays (Costar) for 2 hours at 37°C. PCR screening of recombinants was performed as described above using M13F and M13R primers. Clones containing non-recombinant plasmids gave PCR products of 225 bp, those bearing fragments of the shotgunned insert gave larger products.

### **2.3.14 DNA sequencing**

Only cloned DNA was sequenced.

### **Radioactive sequencing of plasmid DNA**

This was performed using the Sequenase™ Version 2.0 modified T7 polymerase kit (USB) following the manufacturer's instructions. However, 2pmol of primer were used with 2µg double stranded template. Routinely, 5µCi [ $\alpha$  <sup>35</sup>S]dATP were used per reaction. The sequencing primers were derived from the polylinker sequences of the cloning vectors; for pCR II plasmids, the primers pCRSR and pCRSF were used; for pBluescript derived plasmids, the primers KS and SK were used. Sequencing reactions were electrophoresed through 6% (w/v) denaturing (7M urea) polyacrylamide gels in TBE. Subsequently, the gels were fixed in 5% (v/v) methanol 5% (v/v) acetic acid and dried for autoradiography using Kodak™ BIOMAX-MR X-ray film at room temperature for 24 - 48 hours.

### **Automated sequencing**

This was performed with an ABI PRISM™ 377 DNA Sequencer (Perkin Elmer) and in every case, the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit was used (cat. no. 402080). The manufacturer's instructions were observed at all times. Data were analysed with a Macintosh™ Power PC computer using ABI PRISM™ Sequencing v.2.1.1 software.

### **Template preparation for automated sequencing**

Plasmid templates derived from large scale plasmid preparations (after ultracentrifugation) which had been resuspended in water were used without further treatment. Those derived from small scale preparations were resuspended in 50µl TE RNase as usual but were then incubated at 42°C for 30 minutes prior to precipitation with PEG 8000. Washed pellets were resuspended in water at 500ng/µl. In the case of large scale sequencing of differential display products, unpurified plasmid clones gridded in 96 well culture dishes were sequenced after PCR with M13F and M13R in 40µl reaction volumes as described above. 5µl each PCR reaction was visualised after agarose gel electrophoresis, and if a single reaction product was seen, the remainder of the reaction mixture was made up to 100µl and precipitated with PEG 8000. Washed pellets were resuspended in 8µl water and the whole amount sequenced. Differential display clone amplicons (in pCR II and pTAX) were sequenced with the T7 primer only. Shotgun clone amplicons (in pBluescript) were sequenced with both the T7 and the T3 primer.

### **Analysis of sequence data**

Sequence data were edited with a Macintosh™ Power PC computer using DNASTar software (DNASTar). All DNA sequences were interrogated using the BLAST (Basic Logical Alignment Search Tool) set of programmes against databases consisting of

all known genes and expressed sequence tags (ESTs) at the National Centre for Biotechnology Information, National Library of Medicine, Bethesda, MD USA via the world-wide web (<http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST>). A more detailed analysis for weaker DNA homologies was performed using the FASTA software (Pearson WR, Lipman DJ Proc. Natl. Acad. Sci. U.S.A. 85: 2444-2448 (1988)) at the European Bioinformatics Institute, Hinxton Hall, Cambridge, UK also via the world wide web (<http://www.ebi.ac.uk/htbin/fasta.py?request>).

### **2.3.15 Radioactive *in situ* hybridisation**

#### **Preparation of sections for *in situ* hybridisation**

This was performed as described in detail elsewhere (Wilkinson and Green, 1990). Dissected embryos were placed in 10ml ice cold 4% (w/v) paraformaldehyde in PBS at 4°C overnight. They were successively dehydrated through the following solution changes at 4°C for 30-40 minutes each; saline twice, 1:1 saline:ethanol mix, 70% (v/v) ethanol twice, 85% (v/v) ethanol, 95%(v/v) ethanol and absolute ethanol twice. Then the solution was replaced with 10ml toluene twice for 30minutes each, 1:1 toluene:wax mix at 60°C for 20 minutes followed by three changes of wax for no more than 20 minutes each. The embryos were orientated in sectioning blocks before being sectioned at 7µm and applied to Superfrost™ TESP A coated slides (BDH). Sectioned slides were stored at 4°C until required.

#### **Pretreatment and hybridisation of sections**

Slides were dewaxed in HistoClear twice for 10 minutes and after a 3 minute ethanol wash to remove most of the HistoClear, were rehydrated through an ethanol series (absolute ethanol twice, 75% (v/v), 50% (v/v), 25% (v/v) ethanol) followed by saline and PBS for 5 minutes each. They were then washed in 4% (w/v) paraformaldehyde in PBS for 20 minutes followed by PBS for 10 minutes with one change. After this, they were treated with 20µg/ml Proteinase K (Boehringer) in 5xTE followed by a wash with PBS (5 min) and further fixation with 4% (w/v) paraformaldehyde in PBS. After treatment for 10min with acetic anhydride in triethanolamine (0.63ml in 250ml of a 0.1M solution, pH8) the slides were dehydrated through PBS, saline and an ethanol series as for the rehydration step earlier. After the second ethanol wash, the slides were air-dried and either stored at -20°C in an air and watertight container or used for hybridisation by the application of 2.5µl of probe in hybridisation mix per cm<sup>2</sup> clean coverslip. Newly hybridised slides were placed horizontally in a tightly sealed box together with tissue paper soaked in 50% deionised formamide, 5x SSC and incubated overnight at 55°C.

### **Washing of sections and autoradiography**

Coverslips were removed from the slides by submergence in 5x SSC, 10mM DTT at 50°C for 30 minutes and they were washed at high stringency in 50% (w/v) formamide, 2x SSC, 20mM DTT for 30 minutes. After three 10 minute washes in NTE buffer the slides were treated with 20µg/ml Ribonuclease A (Boehringer) in NTE at 37°C for 30 minutes. A further high stringency wash was then performed after a 15 min. wash in NTE alone. After final washes for 15 minutes each in 2x SSC and then 0.1x SSC, the slides were quickly dehydrated through an ethanol series, each of which contained 0.3M ammonium acetate. After the second 100% ethanol wash, the slides were air-dried and used for autoradiography using Ilford<sup>TM</sup> K.5 Nuclear Research Emulsion (Ilford) in 1% (w/v) glycerol. Exposure was for 7-10 days at 4°C in light-tight containers.

### **Development and staining of slides**

The containers with the slides were allowed to warm up to room temperature (2-3 hours) and, under safelight conditions, the slides were transferred through Kodak D19 developer for 2 minutes, 1% (v/v) acetic acid for 1 minute and 30% (w/v) sodium thiosulphate for 2 minutes. They were then transferred to distilled water and counterstained for 3 minutes at room temperature in 3% (w/v) Toluidine Blue. After 1 minute in distilled water, they were dehydrated through an ethanol series culminating in 2 washes in 100% ethanol. They were then transferred into xylene prior to mounting with clean coverslips. DPX mountant (BDH) was routinely used (for which xylene is the solvent).

### **2.3.16 Whole mount *in situ* hybridisation (Long protocol)**

#### **Preparation of digoxigenin - labelled riboprobes**

20µg plasmid bearing the clone of interest was digested with an appropriate restriction enzyme and subsequently purified by phenol extraction and ethanol precipitation followed by resuspension in TE at 1µg/µl. Probes were synthesised using 1µg template in a 40µl reaction volume containing digoxigenin - RNA labelling mix (1x), 10U of the appropriate RNA polymerase (SP6, T3 or T7, Boehringer), 25U Rnasein<sup>TM</sup> (human placental Rnase inhibitor, Promega) and transcription buffer. Reactions were incubated for 2 hours at 37°C and subsequently they were treated with 1U RQ 1<sup>TM</sup> RNase free DNase (Promega) for 10 minutes before fractionation on a Sephadex G50 column set up in a 1ml sterile disposable syringe with a 300µl elution volume in column wash buffer. The column eluates were precipitated with ethanol and the precipitated riboprobe/tRNA pellets were washed

twice with 80% (v/v) ethanol and once with 100% ethanol before resuspension in 100µl TE. Probes were stored at -20°C until required.

### **Preparation of embryos**

Embryos were dissected in ice cold PBS and fixed overnight in a 4% (w/v) solution of paraformaldehyde in PBS. Subsequently, they were dehydrated through 25% (v/v), 50% (v/v) and 75% (v/v) methanol in PBT before immersion in two washes of 100% methanol for 20 minutes each. They could then be stored at -20°C long term until required.

### **Preparation of mouse embryo powder**

12.5 - 14.5 dpc mouse embryos were homogenised with their yolk sacs in a minimal volume ice cold PBS. 4 volumes ice cold acetone were added and after vortexing the mixture was incubated on ice for 30 minutes. It was then spun at 10000g for 10 minutes at 4°C and the supernatant discarded. The pellet was washed again in acetone and respun. It was then spread out on filter paper and ground out into a fine paste before being allowed to air dry. The powder was stored at 4°C in air-tight conditions.

### **Preparation of antibody mix**

100µl embryo powder were suspended in 1ml TBST for 30 minutes at 70°C. The mixture was spun and the supernatant discarded. The pellet was cooled on ice and resuspended in 1ml 1% (w/v) sheep serum. 1µg anti - digoxigenin Fab solution was then added and the suspension rotated at 4°C for an hour before centrifugation at 4°C for 5 minutes at 8000g. The supernatant was mixed with a further 1ml 1% (w/v) sheep serum and stored for up to 1 week at 4°C before use.

### **Pretreatments**

The embryos were rehydrated by taking them through the methanol/PBT series in reverse, washed 3 times for 10 minutes each in PBT at room temperature and bleached using 6% (v/v) hydrogen peroxide in PBT prior to treatment for 15 minutes in 10µg/ml proteinase K in PBT. Then they were washed with fresh glycine 2mg/ml in PBT for 5 minutes at room temperature. After two 5 minute washes with PBT, they were refixed using 0.2% (v/v) glutaraldehyde and 4% (w/v) paraformaldehyde in PBT for 20 minutes and briefly washed again in PBT before prehybridisation in 1ml prehybridisation mix for 1 hour at 70°C. The digoxigenin labelled riboprobe to be hybridised was denatured by heating to 80°C and snap - cooling on ice. It was mixed with 1ml prehybridisation mix and this was added to the prehybridising embryos which were then incubated at 70°C overnight.

## **Washing**

The following morning, the embryos were washed twice in wash solution I at 70°C for 30 minutes, once in wash solution II for 10 minutes at 70°C and three times in solution II for 5 minutes at room temperature. Following this, they were incubated twice in 100µg/ml RNase A in solution II for 30 minutes at 37°C. This treatment was followed by a single wash in wash solution II and then in wash solution III for 5 minutes each at room temperature and subsequently two washes in wash solution III for 30 minutes each at 65°C. After 3 washes in TBST at room temperature, they were preblocked with 10% (w/v) sheep serum in TBST for 2 1/2 hours at room temperature with gentle agitation. The blocking serum was then removed and 300µl antibody mix were added and the embryos washed for 5 minutes. The mix was then removed and 700µl antibody mix were added and the embryos gently rocked overnight.

## **Staining**

The following morning, the embryos were washed three times with TBST for 5 minutes and then received 5 washes in TBST of 2 hours' duration all at room temperature. They were then left overnight in TBST with gentle agitation. After this, they were washed 3 times in NTMT for 10 minutes at room temperature and were then incubated in staining solution comprising 4.5µg/ml NBT and 1.75µg/ml BCIP in NTMT. The staining reaction was allowed to proceed in the dark at room temperature with occasional inspections. When it was deemed to have completed the embryos were washed twice in NTMT at room temperature for 10 minutes and the reaction stopped by a single wash in 1x PBT pH 5.5. The embryos were left overnight at 4°C.

## **Post fixing**

This was done with a 1 hour wash at room temperature in 4% (w/v) paraformaldehyde, 0.1% (v/v) glutaraldehyde in PBT at room temperature. After 2 washes in 0.4% (w/v) paraformaldehyde in PBT they were stored at 4°C in this solution.

### **2.3.17 Whole mount *in situ* hybridisation (Rapid protocol)**

This protocol was developed by Dr. Domingos Henriqué (ICRF Developmental Biology Unit, Oxford).

## **Hybridisation**

The preparation of the probe, embryo dissections fixation and pretreatments were exactly as for the long protocol except that the embryos were prehybridised by rinsing quickly at room temperature in 1:1 PBT/hybridisation buffer C, followed by hybridisation buffer C alone and then were incubated in hybridisation buffer C at 70°C for 1 hour. Overnight hybridisation was then performed at 70°C in hybridisation buffer C containing heat denatured probe.

## **Washing**

The embryos were rinsed twice at 70°C in pre-warmed hybridisation buffer C and then washed twice in it for 30 minutes each, again at 70°C. After a further 20 minute wash at 70°C in pre-warmed 1:1 TBST/hybridisation buffer C, they were rinsed twice at room temperature in TBST followed by two 30 minute washes in TBST, again at room temperature. Subsequently, they were rinsed twice in MABT at room temperature and preblocked with 10% (w/v) sheep serum in MABT for 2.5 hours at room temperature with gentle agitation. The preblocking solution was removed and the embryos incubated overnight at 4°C in fresh MABT containing 10% (w/v) sheep serum and a 1/2000 dilution of the anti-digoxigenin Fab solution supplied by Boehringer-Mannheim.

## **Staining**

The following morning, the embryos were washed three times with MABT for 5 minutes and then received 5 washes in MABT of 2 hours duration all at room temperature. They were then left overnight in MABT with gentle agitation. After this, they were washed 3 times in NTMT for 10 minutes at room temperature and were then incubated in staining solution comprising 4.5µg/ml NBT and 3.5µg/ml BCIP in NTMT. The staining reaction was allowed to proceed in the dark at room temperature with occasional inspections. When it was deemed to have completed the embryos were washed twice in NTMT at room temperature for 10 minutes and the reaction stopped by a single wash in 1x PBT pH 5.5. The embryos were left overnight at 4°C.

## **Post fixing**

This was done with a 1 hour wash at room temperature in 4% (w/v) paraformaldehyde, 0.1% (v/v) glutaraldehyde in PBT at room temperature. After 2 washes in 0.4% (w/v) paraformaldehyde in PBT they were stored at 4°C in this solution.

## CHAPTER 3

### **The initial workup of mRNA differential display.**

This chapter is concerned with the technical development of a differential display system for ultimate use in the context of dissected mouse somites and presomitic mesoderm. Issues involving the reliable extraction and poly (A)<sup>+</sup> enrichment of starting RNA, the reproducibility of differential display with minute amounts of material and efficient cloning of differential display products will be discussed. The initial findings of a comparison of whole mouse embryos at different developmental stages will be presented.

#### **3.1 RNA preparation**

A standard technique for the extraction of RNA from small amounts of tissue is the acid-guanidinium-phenol method (Chomczynski and Saatchi, 1987). It has been widely used in mouse embryology (Harrison *et al.*, 1995) and in situations where RNA is being extracted from only a few cells (Brady and Iscove, 1993; Dulac and Axel, 1995). The method was adopted for these reasons. Other techniques were assessed for example the direct extraction of poly (A)<sup>+</sup> RNA from cell lysates using oligo-dT cellulose (Sambrook *et al.*, 1989) but these were rejected as being unreliable.

An important disadvantage of the method is that with very small volumes, some genomic DNA contamination is experienced. Given that the screening technique was to be PCR based, it was decided to routinely poly (A)<sup>+</sup> enrich all RNA prior to differential display. This had the added advantage of removing ribosomal RNA from the starting material. Oligo-dT coated latex beads (Oligotex-dT<sup>TM</sup>, Qiagen) were chosen for this purpose because the protocol required a single RNA pipetting step and could provide enriched RNA in volumes as low as 20µl. In general, total RNA extracted from small amounts of tissue (*e.g.* dissected somites) was not precipitated but used directly for poly (A)<sup>+</sup> enrichment. In control experiments, an estimated 3.5µg enriched RNA could be extracted from 100µg total RNA using this method.

## **3.2. Differential display**

### **3.2.1 Reagents**

Initially, the conventional protocol of Liang and Pardee, 1992 was followed (refer to section 2.3.7). However one base anchored primers were used for reverse transcription (Bauer *et al.*, 1993; Liang *et al.*, 1994) and [ $\alpha^{33}\text{P}$ ]dATP was used for the PCR step (Bauer *et al.*, 1993; Liang *et al.*, 1995). The steps in the method are itemised below;

1. Reverse transcription was with a single base anchored dT<sub>12</sub>NX primer.
2. One tenth of the reverse transcription reaction was used per PCR.
3. PCR was performed with the reverse transcription primer plus any one of a panel of arbitrarily chosen 10mer primers.
4. The PCR annealing temperature was 40°C throughout

A panel of arbitrarily chosen 10-mer oligonucleotides named O1 to O8 was generated by the NIMR Oligonucleotide Synthesis Service (refer to section 2.1.3).

### **3.2.2 Gel type**

As discussed earlier, it has been well established that native polyacrylamide gels give equally good resolution of PCR products with similar numbers of band differences between starting materials when compared with denaturing gels (Bauer *et al.*, 1993; Liang *et al.*, 1995). Native gels were used in view of this and their relative ease of handling. Best results were to be had if the gels were pre - run for 30 minutes at 4°C and then run at 30V/cm for the desired length of time. The clearest autoradiographic results were achieved with Kodak™ BIOMAX-MR X-ray film. Kodak™ X-OMAT AR and Fuji™ RX were also tried but generally gave inferior results.

### **3.2.3 The effect of poly (A)<sup>+</sup> selection**

Major concerns in differential display centre around the theoretical risks of amplifying contaminant genomic or plasmid DNA species and reverse transcribed ribosomal RNA. It is possible that amplicons from either of these materials could be present as differential bands after the PCR. In order to address these issues, differential display PCR reactions were performed using total and poly (A)<sup>+</sup> enriched RNA from 13.5 dpc mouse embryos before and after reverse transcription.

Figure 2.1 shows the fingerprints obtained when 5µg total and 500ng poly (A)<sup>+</sup> enriched RNA were used as templates for an arbitrarily primed PCR reaction of 40 cycles using O7 without prior reverse transcription. Clearly products are seen in both tracks but are much reduced after poly (A)<sup>+</sup> enrichment. This indicates that (presumably) genomic DNA contamination in the total RNA preparation is largely removed by poly (A)<sup>+</sup> selection. It is of course possible that the total RNA preparation used here was contaminated with extraneous plasmid DNA from elsewhere in the laboratory and that this was eliminated by the enrichment process.

Figure 2.2 shows the fingerprints obtained for each of these templates after reverse transcription. As per the Liang and Pardee protocol, only one tenth of the first strand cDNA was used in the PCR reaction. The amount of starting material indicated above each track corresponds to the amount that was reverse transcribed. A radically different pattern is seen for each with many more bands and a higher general background smear in each lane. This latter feature is probably because the reverse transcription primer is used in the PCR reaction and so, in principle at least, every RNA that is reverse transcribed is amplified at least asymmetrically (*i.e.* with one primer only) 40 times. Although the tracks share many common bands, poly (A)<sup>+</sup> enrichment definitely removes species that are present in the total population. Presumably, these represent ribosomal or other non - poly (A)<sup>+</sup> RNA derived cDNA species.

It is not clear that the differential bands seen in figure 2.1 contribute at all to the patterns seen in figure 2.2. It is possible that the huge excess of cDNA template generated after reverse transcription reduces the probability of significant amplification of whatever small DNA contamination existed in the RNA preparations. During the preparation of total RNA samples derived from very small amounts of tissue, where the extraction volumes are kept necessarily low, the genomic DNA contamination incurred due to pipetting errors is proportionately higher. For this reason alone, poly (A)<sup>+</sup> selection is probably of benefit.

### **3.2.3 The effect of RNA concentration**

The original differential display protocols utilized several hundred nanograms of RNA (Liang and Pardee, 1992). In order to assess the effect of reducing the amount of RNA used on the differential display fingerprint obtained for a given tissue, serial

### **Figure 2.1**

Autoradiograph representing the reaction products of an arbitrarily primed PCR performed on 5µg total and 500ng poly (A)<sup>+</sup> enriched RNA from 13.5 *dpc* mouse embryos without prior reverse transcription. The primer used was O7 and standard differential display reaction conditions were applied. The products were separated by electrophoresis through a native 5% (w/v) polyacrylamide gel run at 50W for two hours at 4°C. All reaction products must be due to contaminating DNA of whatever origin and these are considerably diminished by poly (A)<sup>+</sup> enrichment.

Arbitrary primer 07

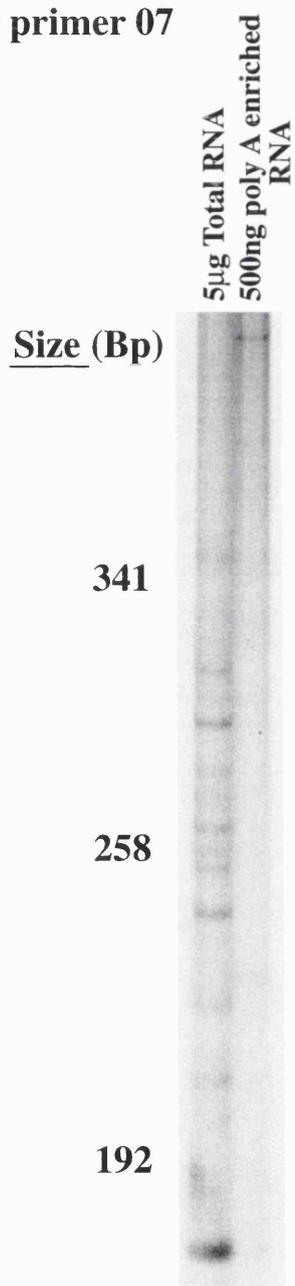


Figure 2.1

### **Figure 2.2**

Autoradiograph representing the reaction products of a differential display PCR performed on 5µg total and 500ng poly (A)<sup>+</sup> enriched RNA from 13.5 dpc mouse embryos after reverse transcription with the primer dT<sub>11</sub>NC. This primer was used along with O7 for the PCR step and standard differential display reaction conditions were applied. The products were separated by electrophoresis through a native 5% (w/v) polyacrylamide gel run at 50W for two hours at 4°C. Although the tracks share many similarities, some species (*e.g.* those marked with the arrows) disappear after poly (A)<sup>+</sup> enrichment.

Reverse Transcription Primer dT<sub>11</sub>NC  
Arbitrary primer 07

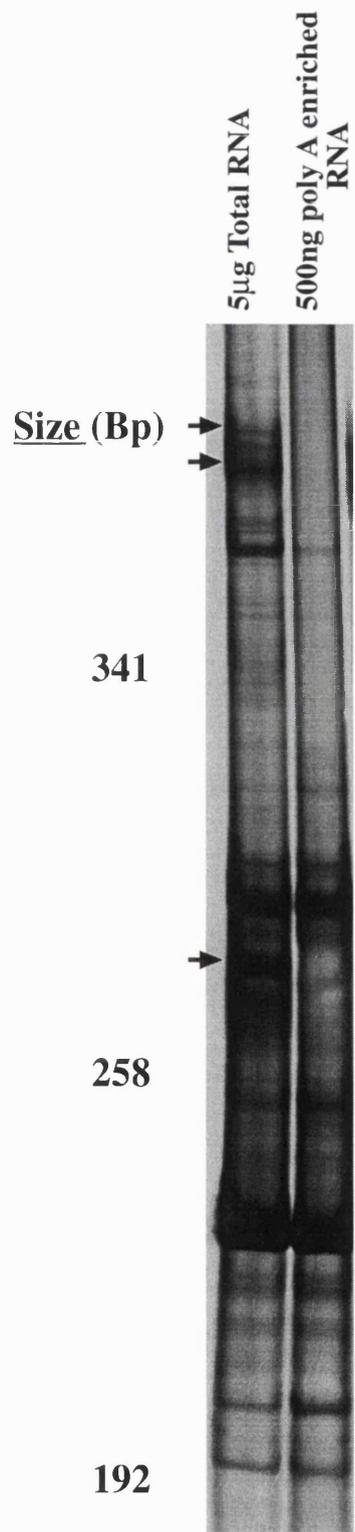


Figure 2.2

dilutions of poly (A)<sup>+</sup> enriched RNA from 13.5 *dpc* embryos were made from a stock concentration of 500ng/μl. 500ng, 50ng, 5ng and 0.5ng aliquots were used for reverse transcription and one tenth of the reverse transcription reaction PCR amplified in a differential display reaction using the arbitrary primer O5. The resulting fingerprints are shown in figure 2.3. The fingerprint pattern obtained is clearly resistant to dilution over this range. Some species *e.g.* band “a”, present with larger amounts of starting material fade at lower concentrations, whilst others *e.g.* band “b”, become more prominent. This rather surprising finding is in agreement with the studies of others (McClelland *et al.*, 1995) who suggested that amplicons that are present after serial dilutions of the starting material are less likely to represent false positive clones. These workers also suggested that performing differential display with less than 50ng starting material is a more reliable way of eliminating false positives than the conventional approach of performing several otherwise identical reactions separately.

During the course of these early studies it was observed that the fingerprints for any single RNA preparation were reproducible for that sample at any given dilution. However, the patterns for RNA samples made on separate occasions from ostensibly the same type of tissue, *e.g.* 13.5 *dpc* embryos using identical primers, whilst similar, were not always identical. This probably reflects the developmental heterogeneity of mouse embryos harvested from the same mother at the same chronological point and also the differences in DNA contamination and the efficiency of poly (A)<sup>+</sup> selection between RNA preparations. These differences were very difficult to formally quantify.

### **3.3. Cloning of differential display products**

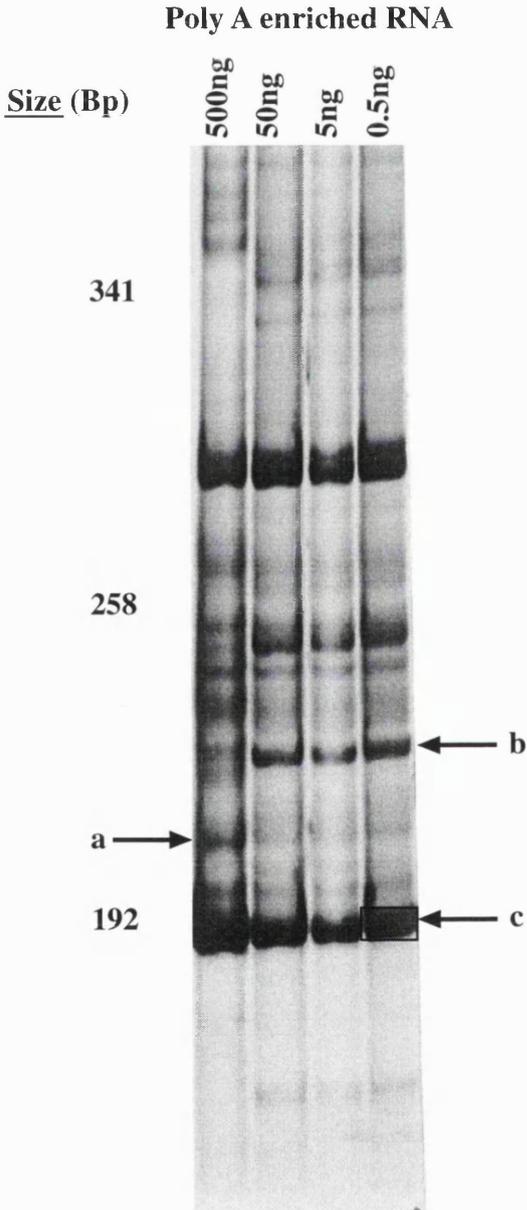
#### **3.3.1 The use of the conventional reamplification protocol**

The use of native gels, which do not contain urea, allowed the direct reamplification of differential display products without the need for precipitation. Typically, excised products were eluted in 100μl distilled water over 1 - 4 hours at room temperature. Initially, a 5μl aliquot of the eluate was reamplified using the differential display primers as described in the original protocol. The efficiency of this second round of PCR was variable, and in general the yield of product was proportional to the intensity of the band excised from the acrylamide gel. The clean excision of a band could be verified by autoradiography and this was done routinely at first. However, once a degree of technical proficiency in the procedure was achieved, it was not performed.

### **Figure 2.3**

Autoradiograph of differential display fingerprints generated from serial dilutions of the same preparation of poly (A)<sup>+</sup> enriched RNA from 13.5 dpc mouse embryos over the range 500ng to 0.5ng. Reverse transcription was with the primer dT<sub>11</sub>NC and the arbitrary primer used in the PCR step in addition to dT<sub>11</sub>NC was O5. Standard differential display reaction conditions were applied and the products were separated by electrophoresis through a native 5% (w/v) polyacrylamide gel run at 50W for two hours at 4°C. Some species, e.g. band “a”, present with larger amounts of starting material fade at lower concentrations, whilst others, e.g. band “b”, become more prominent. The boxed band “c” at the 192bp position was used for the development of subsequent reamplification and cloning technology (*vide infra*).

**Reverse Transcription Primer dT<sub>11</sub>NC  
Arbitrary Primer O5**



**Figure 2.3**

Often fragments larger than 150bp yielded several smaller products after the second PCR and the amount of the correctly sized species in these reactions was insufficient for cloning. This occurred even with discrete bands that could be accurately excised and for which there was no obvious overlap with adjacent species. The phenomenon could occur because the second round PCR reaction was carried out under the same conditions as the differential display and consequently was at low stringency. As has been discussed, because every cDNA species in the PCR reaction can be reamplified at least asymmetrically, then even a gel slice bearing an apparently “clean” band will contain many species of cDNA as well as the species required. In effect a differential display reaction would be performed on this sub-population of cDNA and smaller species derived from it would predominate. Attempts were made to overcome this internal priming by raising the annealing temperature and thereby favouring the reamplification of the desired species. This was not uniformly successful as most primers did not yield any product at annealing temperatures greater than 45°C. If the problem was to be solved, more radical measures had to be adopted as outlined in the next section.

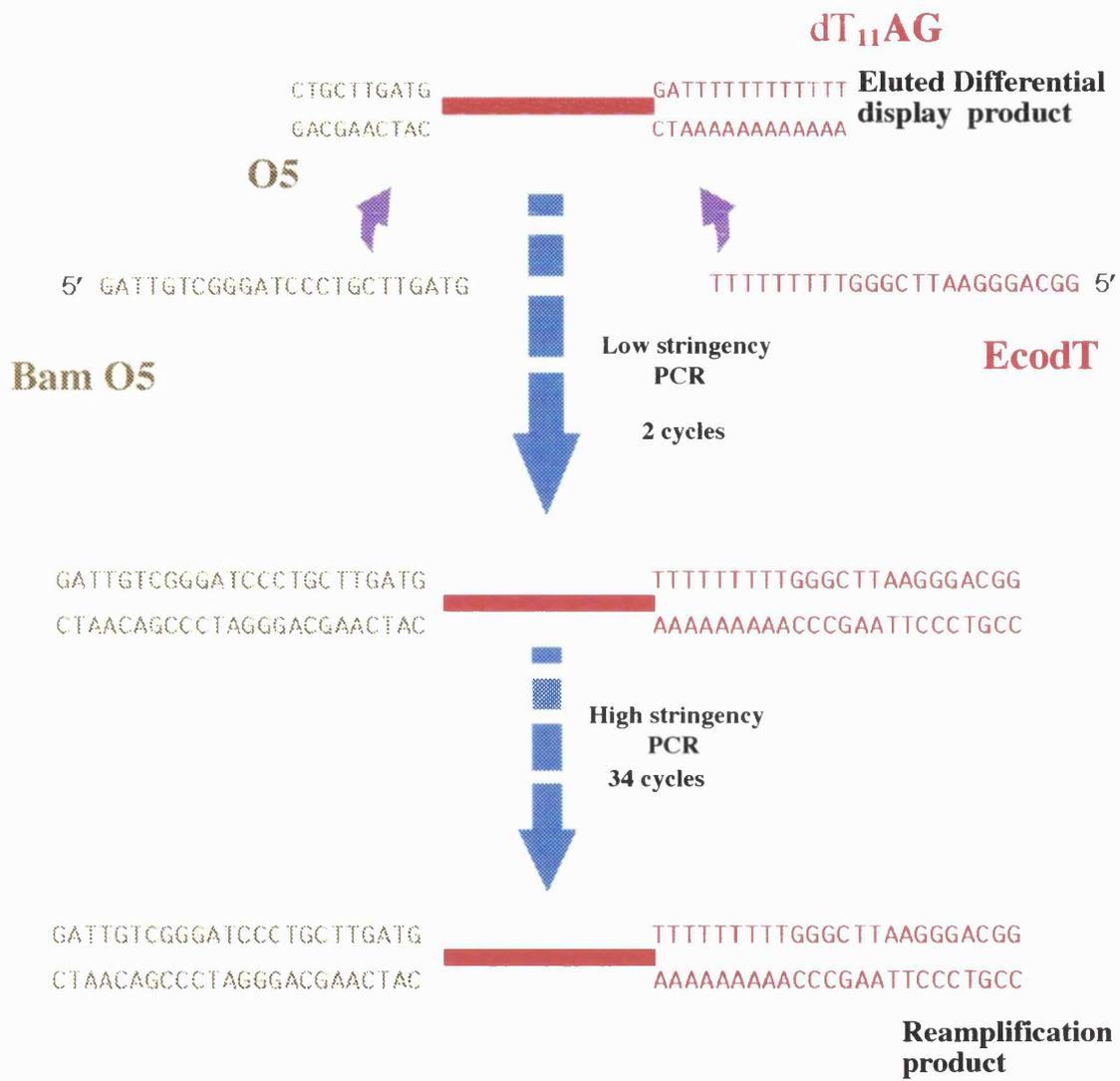
### **3.3.2 A modified reamplification protocol**

In order to increase the yield of reamplified products for cloning and also to ease PCR screening of recombinant clones (*vide infra*), reamplification was carried out with 5' extended PCR primers corresponding to the arbitrary and reverse transcription primers used in the differential display. Figure 2.4 outlines the principle of this approach. For each reverse transcription primer the 24mer EcodT, which is not anchored and bears an *EcoR*I site in its 5' 14-base extension was used. This could prime at the ends of fragments generated by any of the anchored reverse transcription primers. Each arbitrary primer was replaced with a corresponding 24mer carrying the same 5' 14-base extension containing a *Bam*H1 restriction site. After two cycles of PCR with low stringencies of annealing, the products were further amplified for 34 cycles with annealing temperatures of 60°C.

Although this approach, which is equivalent to an AP-PCR reaction performed on the excised gel fragments, does not avoid the problem of internal priming, it increased the yield of all PCR products in the reamplification reaction and thereby facilitated cloning. An example of this is shown in figure 2.5. Here the boxed band “c” in figure 2.3 (size 190 bp) is reamplified either with the appropriate differential display primers (left track) or with the corresponding reamplification primers (right track) and 5µl of

#### **Figure 2.4**

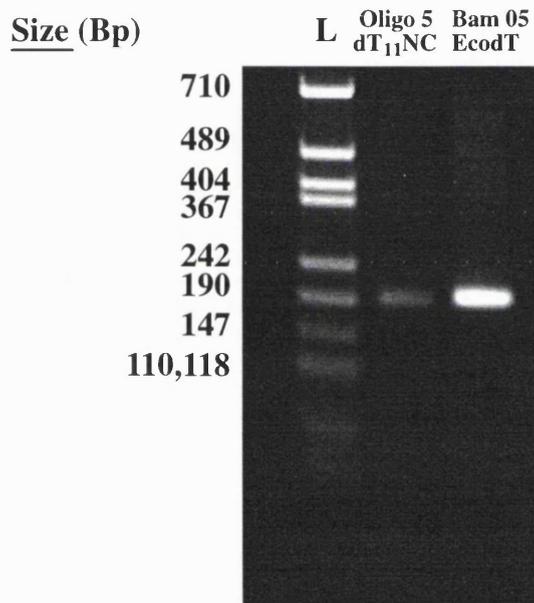
Schematic representation of a modified reamplification protocol for eluted differential display bands that utilises the original differential display primers but with 5' extensions (see text). An idealised double stranded differential display product is shown with the arbitrary primer sequence O5 at one end and the reverse transcription primer dT<sub>11</sub>NC sequence at the other. A universal dT tailed primer EcodT and the specific 5' extended arbitrary primer BamO5 are used for reamplification with two low stringency cycles to allow for incorporation of the reamplification primers into the amplicon followed by 34 cycles at high stringency (annealing at 60°C).



**Figure 2.4**

### **Figure 2.5**

Photograph of the reamplification products of the 190bp fragment band “c” in figure 2.3 after separation by agarose gel electrophoresis. The fragment was eluted in 100  $\mu$ l distilled water and 5  $\mu$ l eluate reamplified with either the original differential display primers (O5 and dT<sub>11</sub>NC) in a 40  $\mu$ l reaction volume using standard differential display reaction parameters or the corresponding reamplification primers (bamO5 and EcodT) in a 40  $\mu$ l volume using the modified reamplification protocol. 5  $\mu$ l each reaction was electrophoresed through the gel (1.5% agarose in TBE). The ladder (L) is provided by the complete digestion of the vector pBluescript<sup>TM</sup> II KS<sup>+</sup> (Stratagene) with *Hpa*II. The yield of product is considerably greater in the latter reaction.



**Figure 2.5**

each reaction loaded onto a 1.5% agarose gel in TBE. The product obtained with the extended primers is predicted to be 23-27 bp longer than the differential display fragment. In this example internal priming is not a problem.

### **3.3.3 Cloning vectors**

The commercially available T/A cloning vector, pCR<sup>TM</sup>II (Invitrogen) or the laboratory generated T/A cloning vector, pTAX (derived from pBluescript<sup>TM</sup> II KS+ (Stratagene) as described) were used for cloning differential display fragments. Both were equally efficient (20 - 30% false positive clones, including cloning artefacts, after *lac* screening) though the latter had a considerable cost advantage because it was made in large quantities in the laboratory.

Reamplified PCR products were not gel purified prior to cloning because this was generally found to reduce the efficiency of ligation and/or transformation. Screening of recombinant clones by PCR provided a rapid way for differentiating which clones bore inserts of the correct size for the band concerned (*vide infra*). The few PCR products refractory to the T/A cloning approach were either blunt-end cloned after end polishing or sticky-end cloned after digestion at the restriction sites in the reamplification primer sequences. Generally, chemical transformation of laboratory stock Epicurian Coli<sup>TM</sup> XL1-Blue or NM522 cells (Stratagene) was sufficient to provide enough clones for analysis.

### **3.3.4 Screening of recombinant colonies**

Initial screening was by conventional *lac* screening and possible recombinant clones were picked into 96-well culture dishes containing selective medium. In order to reduce the incidence of cloning artefacts (*i.e.* clones containing inserted DNA that did not have reamplification primers at their ends), these clones were further screened by PCR using the appropriate reamplification primers. This procedure was made possible by the use of 5' extended primers for reamplification as PCR screening of bacterial clones with short primers at low annealing stringencies was found to be much less reliable. With the arrival of automated sequencing machines at NIMR, the 96 - well tray format allowed for the rapid generation of plasmid derived PCR products from positive clones using the primers M13F and M13R as described in section 2.1.3. Such products could be sequenced directly using internal vector primers.

### **3.4. A comparison of RNA from 7.5 dpc and 8.5 dpc embryos**

In order to further test the reliability of the system using limited amounts of tissue, poly (A)<sup>+</sup> enriched RNA derived from late streak stage (7.5 dpc) and 10 - 16 somite stage (8.5 dpc) embryos was compared. These were chosen as representing embryos with and without somites, and by reverse transcription PCR analysis expression of the somitic marker *myogenin* could be detected in the 8.5 dpc RNA only as shown in figure 2.6. This analysis is not intended to be quantitative, but acts as a rapid quality check for the RNA under study. It was hoped to assess the feasibility of large scale secondary screening by *in situ* hybridisation as northern blots or RNase protection assays would not be routinely possible for the 7.5 dpc embryos.

Differential display analyses of these RNA populations using two different reverse transcription primers and the same panel of arbitrary primers are shown in figure 2.7a. It was not essential to attempt to screen the entire cDNA populations of these tissues, and so a small panel of primers were used. As before, the fingerprinting patterns were reproducible for any given RNA preparation. Poly (A)<sup>+</sup> RNA from the equivalent of fifteen 7.5 dpc and two 8.5 dpc embryos was compared. It was thought that these amounts of tissue from the two stages were of equivalent mass when visualised under a dissecting microscope and consequently that they would give rise to roughly equal amounts of RNA. The areas of the gels shown contain the PCR products that could be most reliably reamplified and cloned. Larger products (up to 600bp) were seen but considerable problematic internal priming was encountered when they were reamplified. Clean excision of the first few bands excised from a gel was verified by autoradiography.

A worrying observation at this early stage was that areas of a gel track that did not contain visible PCR products could be reamplified with the appropriate reamplification primers corresponding to the differential display reaction for that track and the products cloned. This finding adds credence to the notion that in the background of any differential display track which may or may not contain visible bands, there exists a continuum of partially amplified cDNA species.

#### **3.4.1 Sequence analysis**

In general only clones which had inserts of the predicted size for the differential display product were sequenced and analyzed further, although initially even the smaller ones derived from a single band were analyzed. The largest clone obtained

### **Figure 2.6**

Photograph of the reaction products of a qualitative reverse transcription analysis of RNA from pooled late streak stage (7.5 dpc) and 10 - 16 somite (8.5 dpc) stage mouse embryos after separation by agarose gel electrophoresis. Poly (A)<sup>+</sup> enriched RNA from the equivalent of one 8.5 dpc or seven 7.5 dpc embryos were reverse transcribed with a dT<sub>13</sub> primer in a standard 20ml reaction and 2ml of the resulting cDNA was used as a template for a 40 cycle PCR with three primer pairs specific for the genes *hypoxanthine-guanosine phospho - ribosyl transferase (HPRT)*; *myogenin (MGN)* and *oct -3*. The primer pairs were: *HPRT*; HPRT 1A and HPRT 1B (which yield a 354bp product from cDNA and a 575bp product from genomic DNA); *MGN*; MGP1 and MGP2 (which yield a 182bp product from cDNA and genomic DNA); *oct -3*; OC31 and OC32 (which yield a 372bp product from cDNA only). As expected, *HPRT* and *oct-3* are detected in both tissue sources, but *MGN* is amplified only from the 8.5 dpc embryo cDNA. This indicates that the mRNA is free of detectable genomic DNA and that it contains transcripts specific to early mouse embryos with the anticipated differential expression pattern of at least one of these.

7.5 dpc

8.5 dpc

Size (Bp)

L HRT MGN Oct-3 HRT MGN Oct-3

710  
489  
404  
367  
242  
190  
147  
110,118

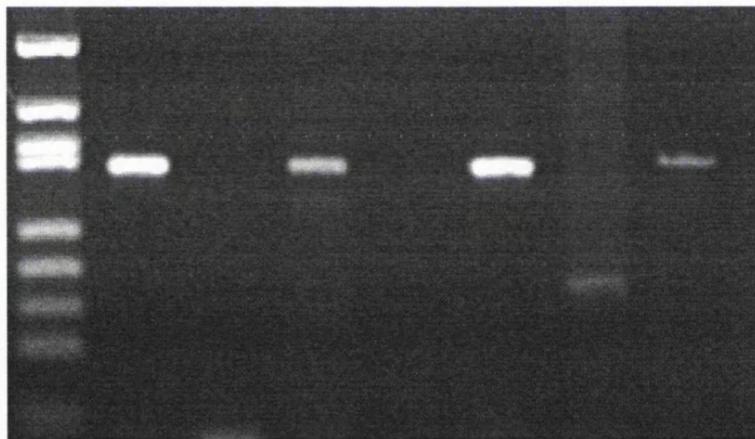


Figure 2.6

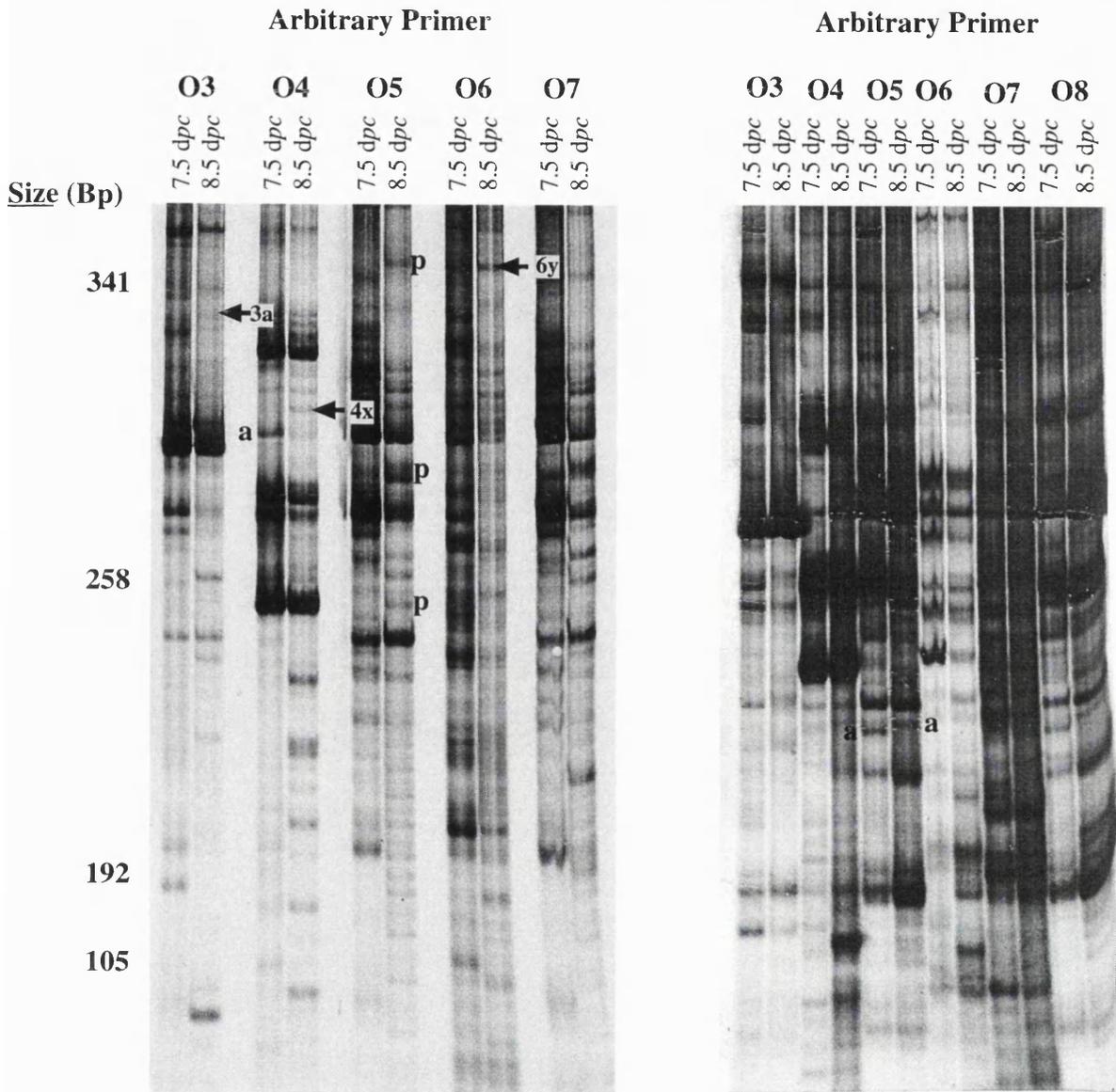
### **Figure 2.7**

(a) Autoradiographs of a differential display analysis of the RNA sources (7.5 *dpc* and 8.5 *dpc*) used in figure 2.6 after reverse transcription with the primers dT<sub>11</sub>NG (left) and dT<sub>11</sub>NC (right). The conventional protocol (Liang and Pardee, 1992) was followed except that radiolabelling was with [ $\alpha^{33}\text{P}$ ]dATP (Amersham) and the reaction products were separated by native polyacrylamide gel electrophoresis. Each adjacent pair of tracks represents the reaction products derived from 7.5 *dpc* (left) and 8.5 *dpc* (right) RNA using the same arbitrary primer as indicated above after reverse transcription with the dT primer used in the gel concerned. The bands marked “a” are those which after reamplification gave clones of a single sequence type. Each of the bands marked “p” gave after reamplification at least one clone that encoded some part of the gene for prolyl oligopeptidase. The bands marked 3a, 4x and 6y gave rise to the clones 3a1, 4x7 and 6y3, respectively, that were used for Northern blot analysis.

(b) Phosphorimages of Northern blot filters made from 1 $\mu\text{g}$  Poly (A)<sup>+</sup> enriched RNA of 8.5 *dpc* (10 -16 somite stage) embryos after denaturing agarose gel electrophoresis that have been hybridised with antisense riboprobes of the clones indicated (The GAPDH (*glyceraldehyde - 3 - phosphate dehydrogenase*) probe was a kind gift of Dr. D. Abrahams, Laboratory of Gene Structure and expression, NIMR). The bands marked 3a, 4x and 6y in figure 2.7 a gave rise to the clones 3a1, 4x7 and 6y3, all of which failed to give a discernible *in situ* hybridisation pattern. Riboprobes were generated in each orientation from 3a1, 4x7 and 6y3. Clear hybridisation signal was seen in only one orientation in each case (shown here).

(a) Reverse Transcription Primer dT<sub>11</sub>NG

Reverse Transcription Primer dT<sub>11</sub>NC



(b)

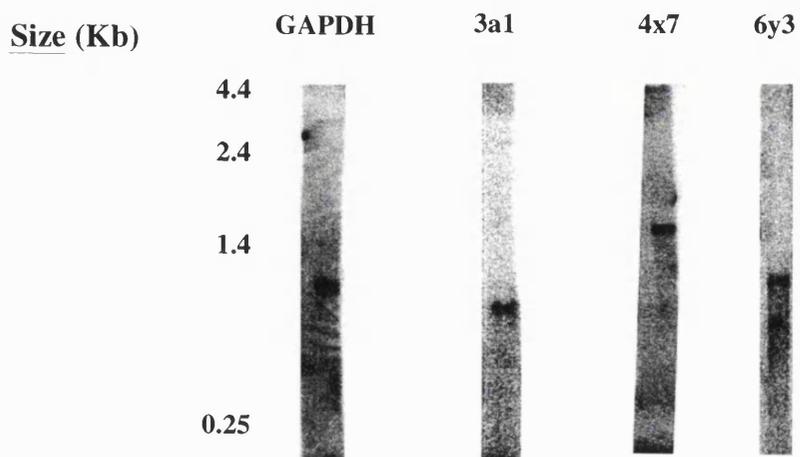


Figure 2.7

was 345bp in size, but the mean size was 206bp and the mode size range was 190 - 200bp.

#### **3.4.1.1 Differential bands**

In this experiment, 203 clones derived from 42 differential bands were studied. They represented 132 individual sequences. 58 had a different primer at each end as predicted by the experimental design. The rest had only one of the two reamplification primer sequences (40 having the arbitrary primer only and 34 having the reverse transcription primer only). This is in keeping with the findings of others (Donohue *et al.*, 1994; Ayala *et al.*, 1995; Diatchenko *et al.*, 1996a).

Only three bands gave rise to clones of only one sequence type and size (marked “a” in figure 2.7a ). Even differential display bands that apparently reamplified cleanly to give a single product could yield up to four clones of identical size but different sequence which could have either or both primer sequences at their ends. There was no obvious way of predicting simply by examining the gel which bands would give rise to clones of multiple sequences or sizes and this occurrence did not appear to be due to errors in triangulation. When smaller clones were generated from a band, they usually included some sequences overlapping with the larger species cloned and these could have any permutation of primer sequences at their termini.

The phenomenon of redundancy (the same cDNA sequence arising from several different differential bands in the same track) was seen only once in this analysis. Clones derived from each of the three bands marked “p “ in figure 2.7a , which was generated from 8.5 *dpc* RNA with the primers dT<sub>11</sub>NG and O5, included the sequence for prolyl oligopeptidase. Each clone corresponded to a different non - overlapping part of the coding sequence and contained only primer sequence corresponding to O5. Seven other unrelated clones were also obtained from these bands. Each of these individual sequence types came from only one of the three bands and each had both primer sequences at its ends.

#### **3.4.1.2 Non - differential bands**

Ten pairs of bands that were of the same size in adjacent tracks were analyzed in order to ascertain that they gave rise to clones with the same sequence. In each case, multiple clones were obtained from each including identically sized species with different sequences. However in every case at least one of these was the same as a

clone derived from the sister band, with identical primer sequences. There was no example of clones obtained from a non-differential band also being isolated from a differential band in the same track.

#### **3.4.1.3 Up- or down- regulated species**

Three pairs of bands that were of identical size but different intensity in adjacent tracks were also studied. These are thought by some workers to represent species of differing abundance between the starting RNAs (Mathieu-Daudé *et al.*, 1996b). As with sister bands of equal intensity, although more than one clone was derived from each species, at least one of these was identical between the two.

#### **3.4.1.4 Primer annealing sites**

No specific primer used here gave rise to a preponderance of clones containing any particular repeat motif, though some used by other workers have been reported to do so (McClelland *et al.*, 1995). In accordance with the findings of others, where a known cDNA was cloned, the 3' most four to six bases of the arbitrary primer were usually exactly homologous to the target sequence (Liang *et al.*, 1992). The sites of priming could be anywhere within the cDNA sequence and not just at the 3' end. In addition, the EcodT primer sequence could lie at the upstream end of the clone. Both of these findings are at variance with the theoretical predictions of differential display, which is supposed to give rise to 3' directed clones because of the dT - based primer initiating DNA synthesis in the second cycle of PCR from the poly (A) rich 3' end of the cDNA to be amplified.

#### **3.4.2 Database interrogation**

As mentioned earlier, database interrogation was performed via the world - wide web using the BLAST (Basic Logical Alignment Search Tool) programs available on the Entrez server at NCBI or the GenomeNet server in Tokyo. All sequences were searched at the DNA level for homology against the entire DNA database including ESTs (BLASTn) as well as at the protein level in the form of conceptual translation against the cumulative protein databases (BLASTx) and a translation of the sequences of the EST database (tBLASTx).

No evidence was seen of mouse intronic, bacterial genomic, cloning vector or ribosomal DNA sequences, thereby justifying the decision to poly (A)<sup>+</sup> enrich the starting templates. Initially, the majority of clones were not homologous to sequences

on the databases, but with the growth of the EST database to its present expanding size, currently 106 of the 132 different clones derived from differential bands have some coding homology with or are identical to known sequences, and of these 52 are represented only in the EST database.

No clone sequenced corresponded to a gene previously known to be differentially expressed between the embryonic time points studied here. In addition the identical sequences derived from sister bands of different intensities did not correspond to any genes known to be down- or up- regulated over this period of development. Differential bands could give rise to clones of genes that were probably not differentially expressed, *e.g.* housekeeping genes. However, no such genes were derived from bands that were non-differential.

### **3.4.3 Secondary screening**

The limited tissue availability from mouse embryos and the wealth of data obtainable per experiment make *in situ* hybridisation the secondary screen of choice for the analysis of the differential display clones generated here. [<sup>35</sup>S]UTP labeled radioactive *in situ* hybridisation studies of slide mounted sections of 7.5 dpc and 8.5 dpc embryos were performed. This technology was routine in the laboratory at the time and allowed for the analysis of multiple planes of section from different embryos on any one slide. The disadvantage of missing expression of low abundance species in discrete anatomical areas not covered by the sections was taken into account, but it was felt at the time that this method would be more tolerant of the small probes used than non - radioactive *in situ* hybridisation analysis. The protocols for both types of *in situ* hybridisation analysis allowed for a maximum of 12 probes to be studied at once.

Of the 132 clones obtained, 30 were discarded as housekeeping or other probably non-differential species. The rest had identity with ESTs, homology with known proteins or were completely unrelated to known sequences. 96 of these were studied by *in situ* hybridisation with riboprobes generated in each orientation. Each riboprobe was applied to two slides bearing sections of 7.5 dpc and 8.5 dpc embryos in various planes. Hybridisation was at 55°C (low stringency) for 15 hours. For each riboprobe studied, one slide was developed after 7 days and the other after 14 days. *oct -3* and *cardiac actin* antisense riboprobes were used as controls.

No clone studied gave a hybridisation pattern consistent with the primary screen, *i.e.* could be detected at one stage but not the other as suggested by the appearance of the

original band in the differential display. Four clones gave ubiquitous hybridisation with each riboprobe in all tissue sections. Each of these clones contained repeat motifs, and it is possible that this contributed to the observed signal with probes in each orientation. 27 showed ubiquitous hybridisation with one riboprobe in all tissue sections and the rest failed to give any hybridisation signal at all despite the low stringency of hybridisation. The smallest probe to give a signal was 235nt but several clones with larger inserts than this failed to hybridize. The five longest clones that gave no signal were subsequently analyzed using whole mount *in situ* hybridisation. No signal was seen with any probe. The short mean insert size of the differential display clones meant that they provided inefficient probes for *in situ* hybridisation; the recommended size range is typically 500 - 1500bp (Wilkinson and Green , 1990), Probe hydrolysis is practised by some workers prior to the hybridisation step in order to improve penetration; however this approach assumes a conventionally sized probe to start with. In embryological systems at least, there is no guarantee that any particular probe generated from a given gene, no matter what its size, will be a useful *in situ* probe. The differential display protocol predicts the generation of clones biased towards the 3' ends of differentially expressed genes. In principle these should generate effective *in situ* hybridisation probes if they are large enough as this region is generally not conserved. The experience in this study was that the clones were not only small but could correspond to any part of the full length cDNA.

Although formal secondary screening by northern blot was not feasible on account of limited tissue availability from 7.5 *dpc* embryos, it was decided at this point to use riboprobes from the three largest putative 8.5 *dpc* embryo clones that gave no signal by *in situ* hybridisation to probe northern blot filters of RNA from 8.5 *dpc* embryos. The sequence of each clone was unrelated to any on the databases at the DNA or the protein level. As seen in figure 2.7b, a riboprobe from each gave a single band by this analysis, implying that is capable of giving a specific RNA hybridisation signal.

### **3.5 Summary**

This preliminary study showed that differential display could generate reproducible reverse transcription PCR fingerprints and that these fingerprints were unchanged over a wide range of RNA concentrations.

Nevertheless there were major problems with the technique. Multiple clones (typically three or four) could be obtained from apparently single differential bands including the isolation of clones with identical insert size but different sequence. However, when products derived from sister (non - differential) bands were cloned, at

least one of the sequences obtained would be common. The prediction of Liang and Pardee (1992) that the clones obtained would be directional and biased towards the 3' ends of genes was clearly not fulfilled. They could arise from any part of the target cDNA and could have any combination of primer sequences at their ends including the occurrence of the reverse transcription primer sequence at the upstream end of the gene. Their small size meant that they generated riboprobes unsuitable for secondary screening by *in situ* hybridisation. Confounding this problem, the method gave no control over where in a gene the clone would reside. Although at least some of them were shown to detect RNA signals at 8.5 dpc after filter hybridisation there was no evidence that any species cloned was truly differentially expressed between the starting materials.

Before proceeding with the method as a way of screening for differential gene expression between dissected mouse somites and presomitic mesoderm, it had to be shown that at least one of the clones that could be isolated from a differential band was indeed differential by some convincing secondary screen. Only then could the issues of eliminating the false positive clones and making the clones large enough for a secondary *in situ* hybridisation screen be addressed.

## CHAPTER 4

### Differential plasmid library screening using arbitrarily primed PCR

The aim of the work described in this section was to establish conditions under which at least one of the clones derived from a differential band was truly differential between the cDNA sources under study. The experience of the first experiments suggested that the easiest way to achieve this was by a filter hybridisation step as a secondary screen. Such a screen would be unfeasible on a large scale if dissected embryonic tissues were used, and so it was decided to apply an arbitrarily primed PCR screen to detect clones differentially expressed between plasmid cDNA libraries made from different murine embryonic tissues. Using this approach, it would be possible to use plasmid library Southern blot hybridisation as a secondary screen for the clones derived from differential bands. Other options were considered at this time, such as studying the embryos of different organisms, *e.g.* chicks or frogs, where embryonic structures are bigger and RNA availability would consequently be less of a problem, but it was decided to persevere with the mouse because the biology of myogenesis and somitogenesis in other vertebrates is clearly different from mammals (Tam and Trainor, 1994).

#### **4.1 Differential plasmid library screening by arbitrarily primed PCR**

The directionally cloned plasmid libraries generated from the germ layers and primitive streaks of gastrulating embryos that have been constructed in the laboratory of Dr. Rosa Beddington at NIMR have proven to be a valuable resource (Harrison *et al.*, 1995). clones differentially represented between the libraries have been isolated by subtraction, differential screening and domain - specific PCR and some of these at least have been shown to be differentially expressed by *in situ* hybridisation analysis (S. Harrison, S. Dunwoodie and P. Thomas; personal communications). The occurrence of clones, generally of low abundance, that are enriched in one particular library but are not differential by *in situ* hybridisation pattern has been observed and this is generally regarded as an artefact of library construction and subsequent amplification. Nevertheless, it was decided to use the libraries for a screen using PCR under differential display conditions. The vector in which they have been constructed (pSPORT1<sup>TM</sup>(GIBCO - BRL)) carries the T7 RNA polymerase promoter upstream of its multiple cloning site, and consequently representative *in vitro* synthesized RNA with a primer derived poly (A)<sup>+</sup> tail can be generated from each library for differential display. Assuming complete linearisation of each template, then the RNA

generated should represent the insert sequences of each library and contain only a small amount of polylinker (vector) sequence. Plasmid Southern blot hybridization could verify whether at least one clone derived from a differential band was differentially enriched between the libraries and if any appeared interesting at the sequence level, then the corresponding full - length species could be isolated for further study.

#### **4.1.1 An outline of the library construction**

This was carried out by Drs. Steve Harrison and Sally Dunwoodie using tissues dissected by Dr. Rosa Beddington in the division of Mammalian Development, NIMR. The libraries were derived from the total RNA of the dissected germ layers of 265 7.5 dpc embryos after reverse transcription with a 5' *NotI* - dT<sub>15</sub> primer. Second strand synthesis was with *E. Coli* DNA polymerase in the presence of *E. Coli* DNA ligase and RNase H, followed by *SalI* adapter ligation (Harrison *et al.*, 1995). The cDNA was cloned into *NotI/SalI* cut pSPORT1<sup>TM</sup> vector (GIBCO - BRL) after digestion with *NotI* and *SalI* and size fractionation. Following electrotransformation and titration, plasmid DNA was isolated by alkaline lysis and density gradient ultracentrifugation from overnight cultures inoculated with bacteria representing known numbers of clones. The mesoderm and endoderm libraries were chosen for study, as they have the largest numbers of clones and biggest mean insert sizes.

#### **4.1.2 Template preparation and differential display**

An equal amount of each amplified library was linearised by digestion with *NotI*, thereby generating a template for representative RNA production using T7 RNA polymerase. It is important in this approach to ensure as complete linearisation as possible and vigorous removal of all DNA. Any unequal contamination of either RNA preparation with vector sequence as either RNA or residual DNA could seriously compromise the validity of the approach. Using 1µg linearised library, RNA synthesis was performed for two hours followed by prolonged DNase 1 treatment. After precipitation with lithium chloride in order to favour the precipitation of RNA over DNA (Sambrook *et al.*, 1989) typically 25µg RNA were produced. 1µg of this was used for differential display in keeping with the original method (Liang and Pardee, 1992).

Figure 3.1 shows the differential display gels obtained when the two RNAs are reverse transcribed with dT<sub>11</sub>NC and PCR is subsequently performed with O5 and

### **Figure 3.1**

Phosphorimages of the differential display reaction products of the *in vitro* translated RNA from mesoderm and endoderm cDNA libraries. The conventional differential display protocol was followed (Liang and Pardee, 1992) except that radiolabelling was with [ $\alpha^{33}\text{P}$ ]dATP (Amersham) and the products were size fractionated by native polyacrylamide gel electrophoresis. For each pair of tracks, mesoderm cDNA is shown on the left and endoderm on the right. The primers dT<sub>11</sub>NC and O5 were used in the first two pairs of tracks firstly without (left) and then with (right) prior reverse transcription with dT<sub>11</sub>NC. The primers dT<sub>11</sub>NC and O6 were used in the final pair of tracks after prior reverse transcription with dT<sub>11</sub>NC. The bands marked were reamplified and the clones obtained used for the secondary screen (figure 3.2).

**Non - reverse  
transcribed  
control**

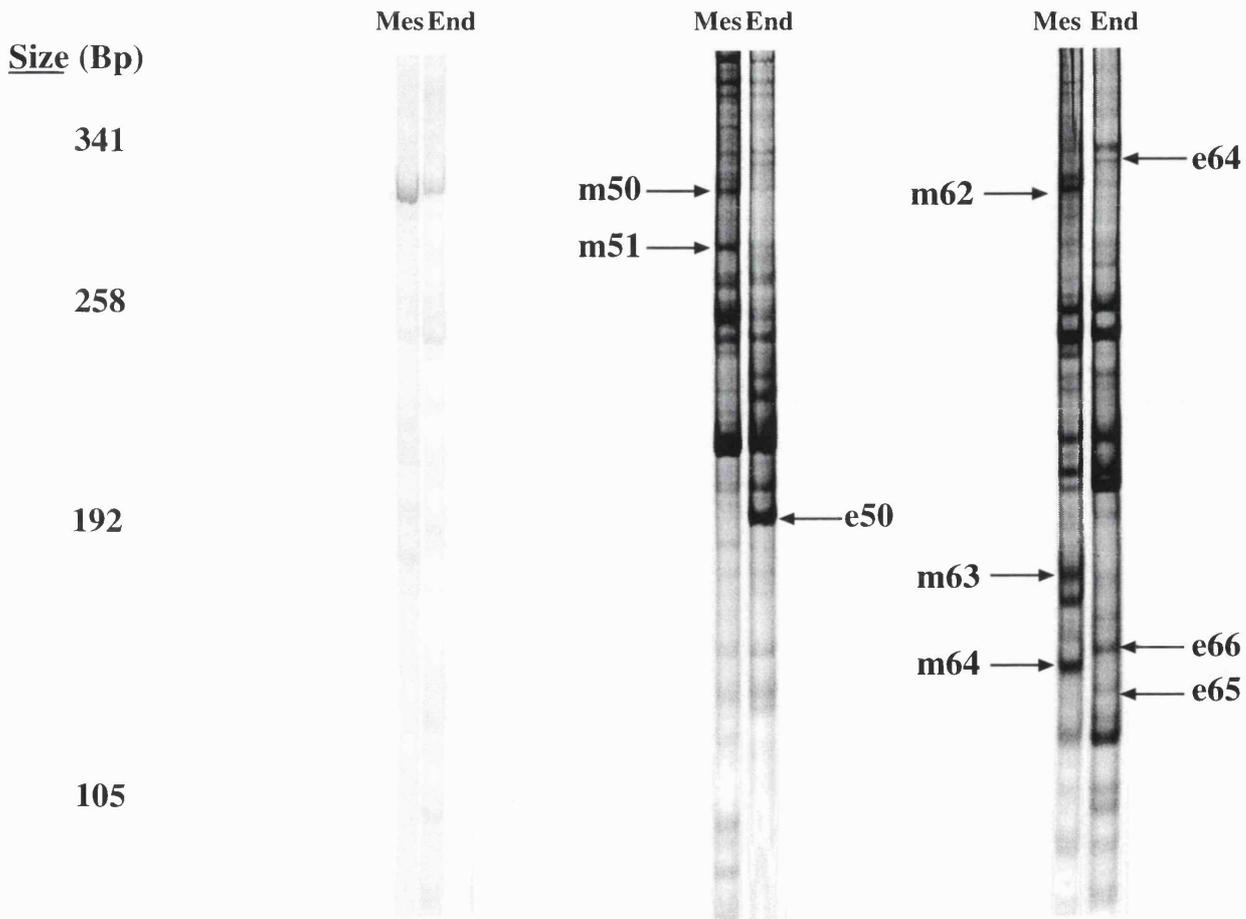
**Reverse transcribed**

**Reverse transcription primer dT<sub>11</sub>NC**

**Arbitrary primers  
dT<sub>11</sub>NC and O5**

**Arbitrary primer  
O5**

**Arbitrary primer  
O6**



**Figure 3.1**

O6. A non - reverse transcribed control experiment is also shown for the primer pair dT<sub>11</sub>NC and O5. Clearly there is some residual DNA contamination but this is minuscule in terms of observed product in the control lanes for each RNA. As with poly (A)<sup>+</sup> RNA generated from tissues, reverse transcription gives rise to a radically different fingerprint with a considerable background smear of labelled products. There appear to be more band differences detected in this experiment in comparison with those carried out on tissue derived RNA. This may indicate differential representation of cloned species in the libraries as a result of the bacterial amplification process.

#### **4.1.3 Sequence analysis of differential bands**

In this experiment, only clearly differential bands were studied and they were eluted and reamplified using the method described earlier. Cloned products generated from each band were picked into 96-well culture dishes containing selective medium and sequenced as before. The size ranges of the products obtained were similar to those from the initial experiments. Of 17 differential bands only one, e50 in figure 3.1, gave a single cloned sequence. From the remainder, between two and five species were isolated per band with similar properties with respect to size, primer content and internal priming as seen in the earlier studies. Evidence of primer redundancy was not seen. Of the sequences obtained, the majority had identity with ESTs from various sources; plasmid or bacterial genomic sequences were not found. No well known mesoderm or endoderm markers were cloned and, as before, the clones were not always 3' directed.

#### **4.1.4 Secondary screening by plasmid Southern blot**

Representative clones from each sequence group derived from a given band were chosen and DNA probe templates were prepared from them by PCR directly from the cultures using the relevant reamplification primers. All templates were gel purified as described and DNA probes were generated by random primer labeling. The clones were given a single letter and three digit identity on the basis of the name assigned to the band from which they were derived, *e.g.* e50 in figure 3.1 followed by the clone number in the row of the 96 - well tray into which they were picked.

For Southern blotting, 1µg each library was digested for 60 minutes with *NotI* and *SaII* and the products separated by 1% (w/v) agarose/TBE gel electrophoresis for two hours at 10V/cm before conventional overnight transfer onto Hybond<sup>TM</sup> - N<sup>+</sup>

positively charged nylon membrane (Amersham). Typically these filters were prepared in bulk and stored at 4°C.

Rapid one hour hybridisation was performed at 60°C in QuickHyb® hybridisation solution as described and hybridisation signals visualised using a phosphorimager (Molecular Dynamics Inc.) and ImageQuant® software on a Macintosh™ Power PC computer.

Examples of this screen are shown in figure 3.2. Figure 3.2a shows the differing hybridisation patterns observed for the two sequence types generated from band e65 in figure 3.1. This should represent an endoderm specific species. One species represented by e654 was abundant in both libraries. The other, e656 hybridised only to the endoderm library. This phenomenon was observed for clones derived from all bands studied. In all 17 cases, at least one clone derived from every band was detectable in one library but not the other or was at least enriched greatly in one over the other.

#### **4.2 Further analysis of differential clones**

Clearly the arbitrarily primed PCR approach adopted here could rapidly isolate from each library clones that were differential. The precautions that had been taken to eliminate vector sequence contamination appeared to have been effective and the large amount of template material that was available meant that secondary screening by filter hybridisation could be employed to identify the clones derived from a given band that were enriched in each library.

In order to further test the validity of this method, it was decided to analyze some of the clones by *in situ* hybridisation. A substantial number of clones isolated from these libraries by differential and subtractive screening approaches that were enriched in one or the other by Southern blot hybridisation did not give *in situ* hybridisation patterns consistent with their representations in the libraries. Often no pattern or ubiquitous expression were observed. However no examples of clones enriched in one library over another by filter hybridisation that gave a reverse representation by *in situ* hybridisation were found (Steve Harrison; personal communication).

##### **4.2.1 Isolation of full-length clones**

As the clones generated from this screen were small, full-length species had to be isolated for the purpose of *in situ* hybridisation. The libraries were plasmid derived

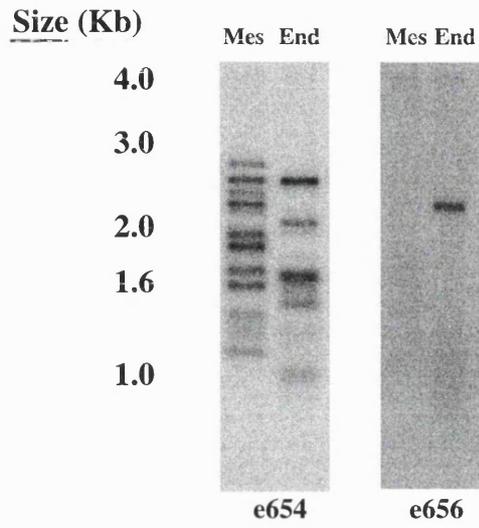
### **Figure 3.2**

Phosphorimages of library plasmid Southern blots hybridised with probes from the clones indicated. 1µg each library was digested with *NotI* and *SaI* for one hour, and the products separated in adjacent tracks through a 1% (w/v) agarose/TBE gel prior to Southern blotting. Probe templates were generated by PCR amplification of the inserts of the clones in question using the appropriate reamplification primers followed by gel purification. Labelling was by random priming.

(a) Clones e654 and e656 represent the two sequence classes derived from the endoderm specific band e65 (figure 3.1). Clone e654 is represented in both libraries but e656 is restricted to the endoderm library.

(b) Filter hybridisation patterns of the clones derived from the endoderm bands e50, e64 and e66 and the mesoderm bands m62, m63, m64, m50 and m51 which were clearly differentially expressed between the two libraries.

(a)



(b)

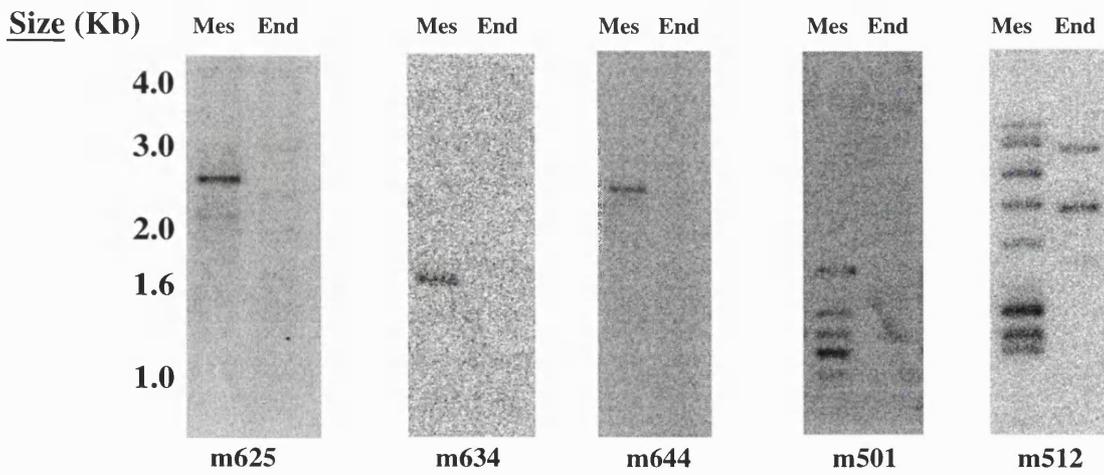
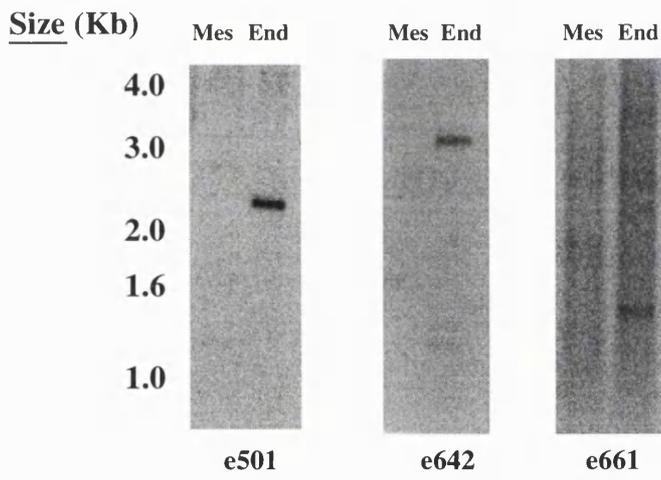


Figure 3.2

and could be screened either in the form of conventional bacterial colony lifts or plasmid Southern blots made from known numbers of clones. The latter method was adopted on account of the high background hybridisation incurred when colony lift filters were hybridized with short DNA probes.

Briefly, the number of clones generated from an aliquot (typically 1 - 5pg) of plasmid library was measured after standard chemical transformation and subsequently enough plasmid was transformed to assure a ten fold over-representation of all the clones in the library. This transformation was aliquoted into 48 pools (with a known number of individual clones) which were grown in liquid culture and plasmid DNA prepared from them. After complete digestion of the DNA from each pool with *NotI* and *SallI*, gel electrophoresis and Southern blotting, filters were hybridized with a probe of the differential display product of interest and pools that hybridized were selected for further sub-pooling by this process. Eventually after three to five rounds of this, positive pools contained fewer than 100 individual clones which could be picked and the inserts amplified by PCR using the primers PE and AH (refer to section 2.1.3). These PCR products were screened by filter hybridisation after gel electrophoresis and Southern blotting. Although the method was cumbersome, it was effective and had the advantage of indicating the insert size of the clone to be isolated. The inserts of the clones isolated by the procedure were used to generate probes for hybridisation screening of the plasmid Southern blots in order to verify the same hybridisation pattern seen with the original differential display clone. The whole process was carried out for eight different clones, all of which were clearly represented in only one of the libraries.

#### **4.2.2 *In situ* hybridisation analysis of full - length clones**

This part of the experiment was performed by Dr. Paul Thomas in the Division of Mammalian Development at NIMR. The conventional long protocol was used with embryos of primitive streak and post-gastrulation stages up to 20 somites. Of those clones analyzed only two (e661 and m634 in figure 3.2) gave *in situ* hybridisation patterns consistent with their representations both in differential display gels and in the libraries. The rest failed to give any convincing hybridisation signal at all or in one case (m625) was expressed at low levels in all three germ layers. Generally, several subclones of various sizes were used to generate riboprobes in such cases. It is possible that the differential display was detecting differences in some extremely low abundance species as suggested by some workers and that their detection by *in situ* hybridisation is beyond the capabilities of the protocol used here (Conway 1995; Guimarães *et al.*, 1995a,b; Rafaeloff, 1996).

### **4.2.3 “Shotgun” cloning for sequence analysis**

A conventional strategy was adopted for this as described earlier. Briefly, plasmid DNA insert released after *NotI* and *SalI* digestion was gel - purified, self ligated and sonicated to a mean insert size of 400 - 600bp. The products were cloned into pBluescript<sup>TM</sup> II KS<sup>+</sup> and recombinant clones amplified directly by PCR using the primers M13F and M13R (refer to section 2.1.3). The PCR products were sequenced directly with the primers T3 and T7. Sequencing of 36 clones derived from a single insert could yield up to 20Kb of sequence from a given clone within three days of its isolation. This typically allowed for a six- or seven-fold redundancy of coverage of the sequences of most clones.

### **4.2.4 Northern blot analysis**

A conventional method was adopted as described earlier (Sambrook *et al.*, 1989). Usually 9.5 *dpc* embryo RNA or RNA derived from a range of adult tissues was used.

### **4.2.5 Clone e661**

This was derived from an endoderm library specific band, e66 in figure 3.1. One other sequence type was obtained from e66 which encoded a mitochondrial gene. The initial reamplified product was 142bp in size and had no sequence homologies at the DNA or protein level with known sequences on any of the public access databases. The secondary screen showed it to be enriched in the endoderm library (figure 3.2 b) and a longer species of 1245bp was isolated from the library using the Southern blot approach. When this was used for *in situ* hybridisation analysis, expression was seen in the extra-embryonic endoderm at the primitive streak stage and in the yolk sac up to the 16 somite stage (figure 3.3a). No older embryos were studied by *in situ* hybridisation and no evidence of intra-embryonic expression was seen. This is a clear example of the differential display producing an endoderm specific band that gave rise to a clone enriched in that library whose *in situ* hybridisation pattern of expression is purely endodermal.

Northern blot analysis revealed a 1.6 - 1.8Kb transcript detectable in 9.5 *dpc* yolk sac and adult testis RNA but not in 9.5 *dpc* embryonic RNA (figure 3.3b). Clones that are enriched in the yolk sacs of early embryos are often expressed later on in the developing liver. Interestingly the message was not detected in embryonic liver RNA

### **Figure 3.3**

Studies of e661.

(a) Photograph of *in situ* hybridisation analysis with an antisense riboprobe derived from a 1245bp subclone of e661. Hybridisation signal is seen in early, mid and late streak stages in the extra embryonic (yolk sac) endoderm. No intra embryonic signal was observed. At later stages upto the 14 - 16 somite stage, the extra embryonic expression was maintained in the fragments of attached yolk sac but no intra embryonic signal was seen.

(b) Phosphorimages of Northern blot filters containing 2µg each of poly (A)<sup>+</sup> enriched RNA from 9.5 *dpc* embryos, 10.5 *dpc* yolk sac and adult testis after hybridisation with antisense riboprobes for GAPDH and e661. A 1.6 - 1.8Kb transcript is detected in RNA from yolk sac and testis but not in RNA from intra embryonic tissues with the e661 probe.

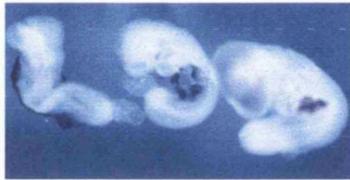
(c) Phosphorimages of Northern blot filters containing approximately 400ng each of poly (A)<sup>+</sup> enriched RNA from 14.5 *dpc* embryonic liver after hybridisation with antisense riboprobes for GAPDH and e6614. Signal is seen with GAPDH only.

(d) Alignment of the e661 protein sequence with those of human and murine BMP 1 and murine *procollagen c proteinase* over the region of the CUB domain. The consensus sequence is shown at the top. Conserved amino acids between the proteins are shown in yellow. Within the CUB domain, absolutely conserved cysteine (C), isoleucine (I), Leucine (L) and glycine (G) residues are highlighted in red; the boxed amino acids represent conserved aromatic (a) and hydrophobic (h) residues.

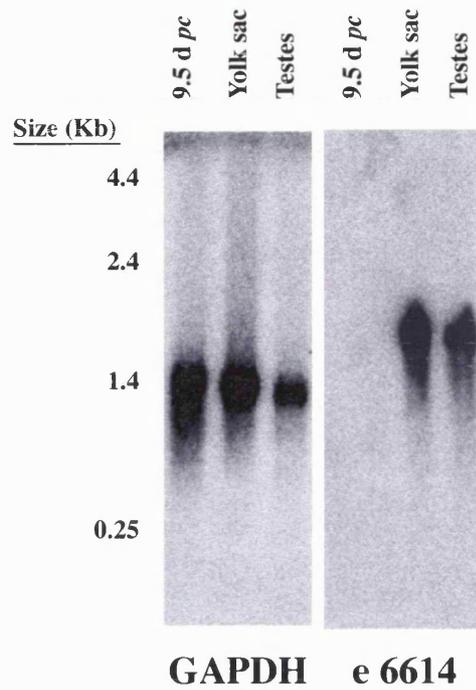
(a) Early, mid and late streak stage



4-6, 8-10 and 14-16 somite stage

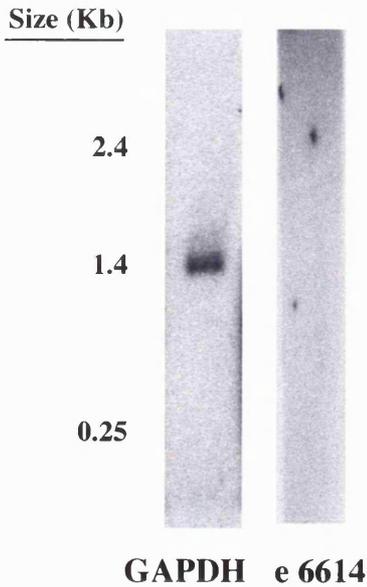


(b)



(c)

14.5 dpc embryonic liver



(d)

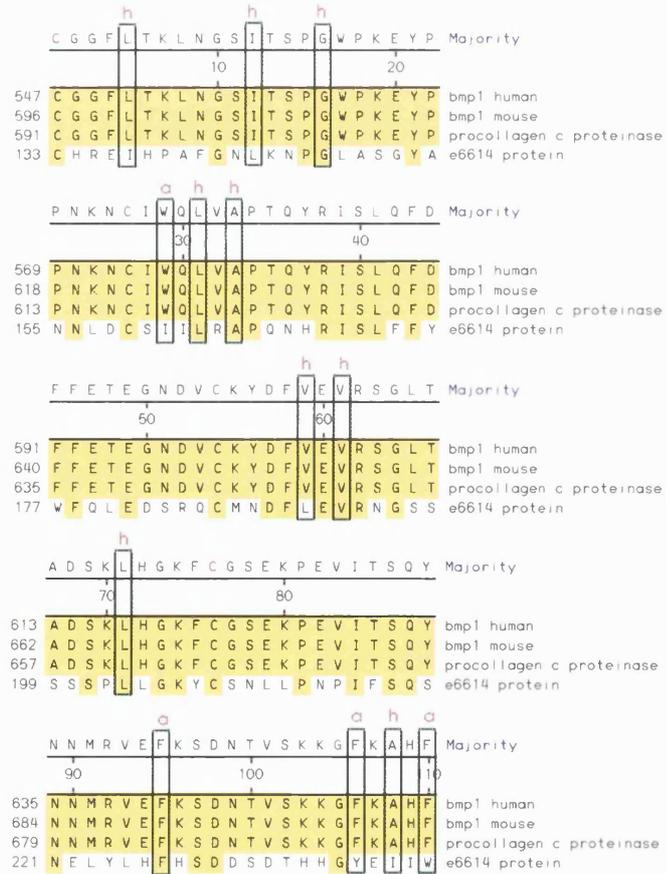


Figure 3.3

isolated from 14.5 *dpc* embryos (figure 3.3c).

A probable full - length clone of 1740bp was isolated after hybridisation to a filter representing 40000 gridded clones of the endoderm library (kind gift of Dr. S. Dunwoodie). Sequence analysis with an average 6 - fold redundancy of coverage revealed an open reading frame of 1263bp giving a conceptually translated protein of 421 amino acids (figure 3.4a, b). Although novel at the DNA level, the sequence was homologous at the protein level to a diverse set of proteins from several organisms including BMP - 1, complement subcomponents C1r and C1s, neuropilin, and tolloid. The region of homology was in a CUB domain (Bork and Beckmann, 1993); the acronym *CUB* is derived from the three proteins in which the domain was first recognized Complement C1r, Sea Urchin embryonic protein and BMP - 1. This is an extracellular motif of approximately 110 - 120 amino acids that is present in a wide range of gene products that are known to play a role in developmental processes, usually as secreted or transmembrane molecules. The protein alignment of this gene with its closest homologues is shown in figure 3.3d. Although an antiparallel  $\beta$  sheet secondary structure and conserved cysteine - cysteine disulphide bridges have been established for CUB domains, there is as yet no information regarding function. No other known motifs were seen in the sequence and there was no clear signal sequence, however a predicted protein kinase C phosphorylation site was observed downstream of the CUB domain (figure 3.4b).

#### **4.2.6 Clone m634**

This was generated after cloning the species derived from the reamplification of the mesoderm generated band m63 (figure 3.1). Three sequence types of 169, 173 and 177bp were isolated, all of which had no homologies at the DNA or protein levels to known sequences. Only the 177bp species represented by m634 was enriched in the mesoderm library (figure 3.2b). This clone was also represented in the ectoderm library (figure 3.5c). Despite considerable efforts, the largest clone that could be isolated from either library was 632bp. Sequence analysis revealed no homologies. *In situ* hybridisation studies failed to reveal expression before the 4 - 6 somite stage when signal could be observed in the dorsal neural tube (figure 3.5 a). By the 16 somite stage expression was seen also in rhombomeres 5 and 7 and also at the midbrain/ forebrain boundary (figure 3.5 a). Transverse sections of these embryos through the mid-thoracic region confirmed expression in the neural tube and also revealed expression in the dorsal root ganglia and possibly adjacent somitic tissue (figure 3.5b). More detailed analyses are currently underway in an attempt to resolve this.

**Figure 3.4**

(a) DNA Sequence of full length 1740bp e661 clone. A putative open reading frame is shown in blue and a consensus polyadenylation signal is underlined in red.

(b) Conceptual translation of the open reading frame shown in figure 3.4a. The CUB domain amino acids are boxed and underlined with absolutely conserved residues in red. A consensus protein kinase C phosphorylation site is shown underlined in green.



### **Figure 3.5**

Studies of m634.

(a) Photographs of *in situ* hybridisation studies with a 632bp anti-sense riboprobe for m634. At the 4 - 6 somite stage expression first appears in the dorsal neural tube (d). By the 9 somite stage, faint expression is also seen in the somites (s). At the 14 somite stage expression is also seen in rhombomeres 5 and 7 (r5, r7) and the forebrain midbrain boundary. The bar across the 14 somite stage embryos corresponds to the line of transverse section used for figure 3.5b.

(b) Transverse section of the 14 somite stage embryo in figure 3.5b. Signal is seen in the most dorsal region of the neural tube (d) as well as the dorsal root ganglia (drg).

(c) Phosphorimages of library plasmid Southern blots hybridised with a m634 DNA probe. The libraries were embryonic stem cell (ES), whole late streak stage embryos (ER), and late streak stage ectoderm (Ect), mesoderm (Mes), endoderm (End) and primitive streak (PS). The filters were made as for figure 3.2 and the probe template was generated by digestion of a m634 plasmid with *NotI* and *SalI* followed by gel purification of the released insert. The clone is represented in the ectoderm and mesoderm libraries only.

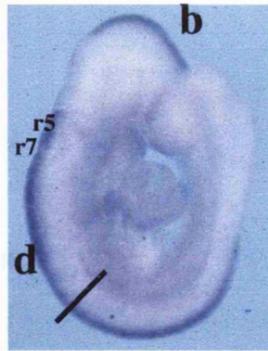
(d) Sequence of the longest m634 cDNA clone isolated. A consensus polyadenylation signal is shown in red.

(e) Phosphorimage of a northern blot containing 5µg poly (A)<sup>+</sup> enriched RNA from 9.5 *dpc* mouse embryos after hybridisation with a 1.384Kb antisense riboprobe for m634. A strong signal is seen at approximately 4Kb.

(a)

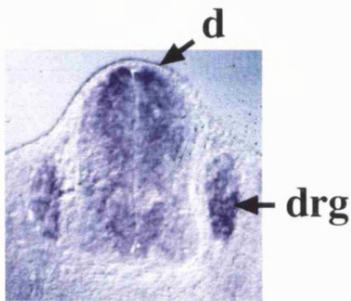


4-6 and 9 somite stages



14 somite stage

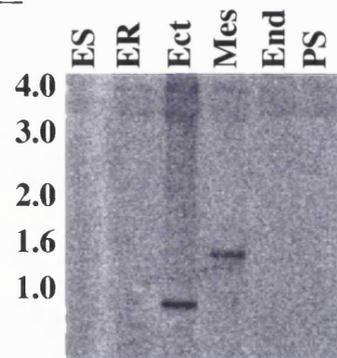
(b)



(c)

Size (Kb)

Library



(d)

```

GGCACGATCACAGTTGGTGGAAAAGGCTGAGGTCCGTGACCTGAAACTCAAAC
TCAAAGGAAAATGACAGCTTTGCCTGCCCTTTGAGGAGTCGAGTTCCAAGACCTCA
CCCCCTTCTCCTGCTAAGGAGCTTCATTTGAGTCCTTCTCCAAGCC TCC TCGGGTCC
TCTCCAGGCTCCTCCTACCCTGCCCTACAGTGAACCTATGGAGCC TACAAGGACC
ATGTTTAAACTAAGGAACACAGAGTCTGGGTGCCAAAAAGTGGAGTCTGGAGA
GCCAACAGCCGAAGCTCCCTTTTGAAGGGTCTTGGGTC TGGACTCCATTAATCT
GGCCCTCCATGACAGCCTTCATGCTCACAGGCCACAGGGAGTATGTGATTACTGT
TAGGCCACGGCAGTTGTGAAGACACATAAAATGGGAGGCCAGTTTAACCGGCTGC
GTTACAGAAAACAAGTGTTTACTGCAGGAATGAAGTTTCCCAAATCTCCAGGAA
GAACAGACATGGGACAGATCCATGGCCAAGCCTCCTCTTGT TTTGCCAAGTCATCT
GGTAAAAATCACAGACCTCCCTTCAAC TTTCCCTAGTCTACATCTGCAGGTACCAA
TAAGACCATCTACATAACCTGCAAGAAACATGTTTAAAAAGCAAAGATCAGGAG
CCCTGGCTTGGCAGGAAATGCGGGTTCTACCCCTTACC TACTGAAGCACAGTCAC
AAGGGTGTCCCTTCCCACTCCAGCGATCTGGCTTCTAATACAATCTCCAGGT
TGCTCTTGGGTGTGATGGGTGGGTGTCACAGGTCCAGGATACAGATGGAAGACA
GATGCTGGGGGCCGGGACAGTCGTTTGGACAGTTCAGC ACTATTGCATTCGCTG
TGGCTTGGCAACTCTGAAGTCTGCTTCC TGGCAACTATCAGCAAGGGCAAAC TTA
ACACCCAGAGCCTGCC TTTGGGCTGTTGGCTGCTGCTTAGAGGACACAAGTGGGG
CCCAGTGGT TAGGAACACTGAATGTTCTCCCAAGGGGACCCATGTTCAATCCAGG
TAGTCTCACAGCTACCTATAACTTCAAGATCTGACACCCTCACCCAGACATAATG
CAGGCAAACATCAATGTACAGGAAATAAAAAATAAACTAACATTTTAGAGGAC
GCCAGGTGGCTCCTAGGCCAGAGCCTTTATGGTTGGACTTTTGGAGTTT TGA
TAATTCCTTGGAGGGCTGTAGTTTTTCCACCTTGG AACGTGGGAGCATCAGGGGA
AGGGGTGTCATGCCTTGATTTATAATGCC TCTTGATGGATTTTCCAAGCTGTA
AATAAGAGAGGTGACAGTTGCTAAGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAA

```

(e)

Size (Kb)

polyA<sup>+</sup> RNA

9.5 dpc

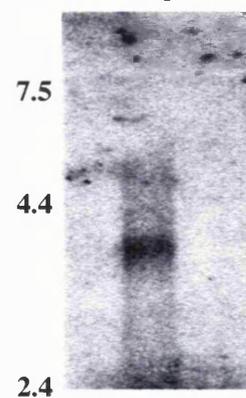


Figure 3.5

Northern blot analysis revealed an approximately 4Kb species detectable in 9.5 dpc embryo RNA (figure 3.5e). No signal was seen in the RNA of any adult tissue tested.

In order to isolate a full length clone, a 12.5 dpc mouse neural tube library that had been directionally cloned into a Lambda ZAP<sup>®</sup> vector (kind gift of Dr. Domingos Henrique, ICRF Developmental Biology Unit, Oxford) was screened. Seven independent clones were isolated, the largest having a 1384bp insert. The consensus sequence of all of these clones is shown in figure 3.5d. It has no homology at the DNA or protein level to any known species. A genomic clone has been isolated and the sequencing project for this gene is currently in progress. Although the m634 clone did not give a detectable signal at the embryonic stage from which the libraries were constructed, it did show some expression in tissues derived from those germ layers in which it was enriched.

#### **4.3 Summary**

This experiment demonstrated that by using filter hybridisation as a secondary screen, differential bands isolated by an arbitrarily primed PCR carried out on representative RNA from the plasmid libraries could give rise to clones representing truly differential species. In principle at least the method had been validated. However, several technical problems remained. In addition to the issues of several clones arising from a single band, their small size not only precluded direct *in situ* hybridisation analysis but also necessitated laborious library screening in order to isolate full length clones for this purpose. Many workers have commented that this stage is often rate limiting in differential display experiments (McClelland *et al.*, 1995). The additional problem of clones being differentially represented in the libraries but not differentially expressed by *in situ* hybridisation was encountered. Allied to this finding is the observation that the differential display in this case tended to generate clones present at low levels in the libraries as assessed by the intensity of the hybridisation. This is in agreement with the general impression of several workers (Conway 1995; Guimarães *et al.*, 1995a,b; Rafaeloff, 1996).

## CHAPTER 5

### **The analysis of differential gene expression in mouse somites and presomitic mesoderm using mRNA differential display**

The work presented thus far has indicated that the differential display protocol can generate reproducible reverse transcription PCR fingerprints over a wide range of RNA concentrations for a given tissue. In addition, differential bands can contain cDNAs that are truly differential between the starting materials by an independent secondary screen. The residual problems that were outstanding at this point were that the clones obtained were small (thereby precluding their direct use for *in situ* hybridisation analysis) and that multiple clones were generated from single bands on the display gels.

#### **5.1 A modified differential display protocol**

As has already been discussed, one of the reasons for the high background of non - differential clones observed in conventional differential display experiments could be that the primer used for reverse transcription is also used in the PCR at a stringency where every species that is reverse transcribed can be amplified, if only asymmetrically. Consequently at any position in the gel, there could exist a large population of species that may or may not be visible but which could be reamplified. This latter possibility is further enhanced by the reamplification conditions also being at low stringency. Even if reamplification is carried out with 5' extended primers, the first two cycles are at low stringency and so, in effect, conditions similar to differential display exist in this reaction. An obvious remedy for this problem would be to somehow remove the reverse transcription primer from the subsequent PCR. When dealing with very small amounts of RNA/cDNA, it is not desirable to attempt this using physical methods because of the risk of losing material. The same effect could be achieved biochemically if the reverse transcription primer could be effectively barred from participation in the PCR on account of a high annealing stringency after the first two cycles.

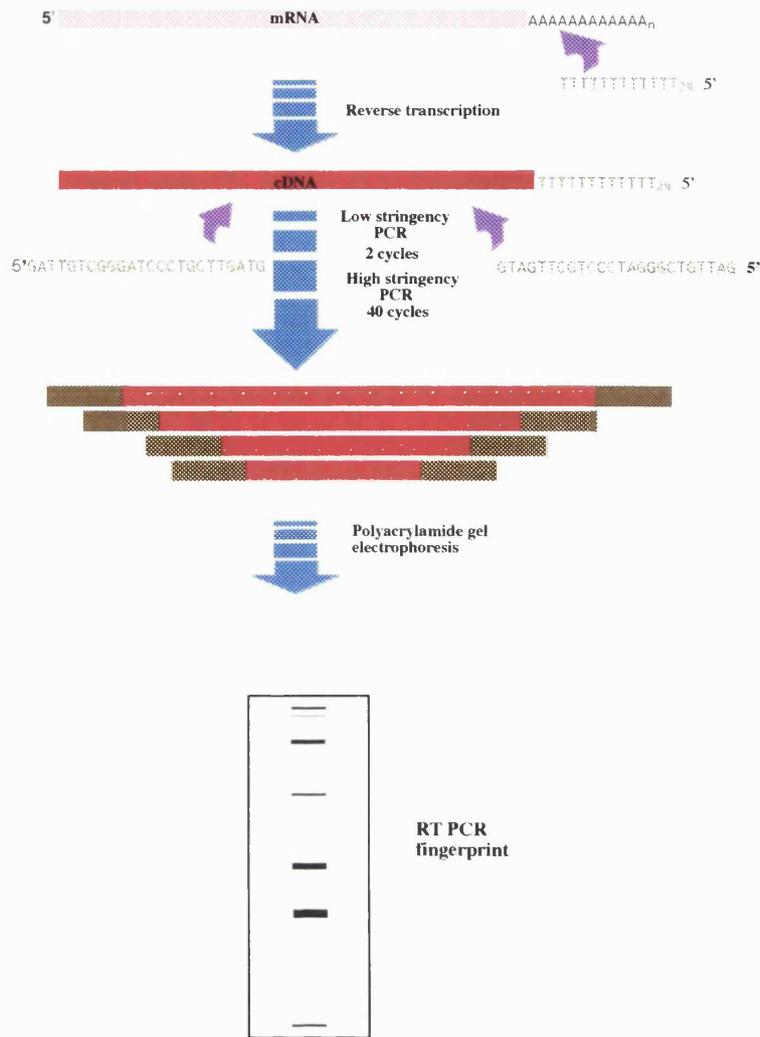
The principle of this idea is shown in figure 4.1a. If reverse transcription is carried out with a conventional dT<sub>24</sub> primer and the subsequent PCR performed with a single arbitrarily chosen 24mer such as those used earlier for reamplification, then after two low stringency PCR cycles that allow the primers to be incorporated into second

### **Figure 4.1**

(a) Schematic representation of a modified RNA fingerprinting protocol designed to prevent the primer used in the reverse transcription step from taking part in the PCR. After reverse transcription with a dT<sub>24</sub> primer, a single arbitrary 24mer is used for PCR with 2 cycles of low stringency followed by 40 cycles at high stringency. This latter step is performed at an annealing temperature at which the dT<sub>24</sub> primer will not bind the template and initiate DNA synthesis, although it may take part in the first two cycles of PCR. The fingerprint observed is derived from the amplification of those cDNAs generated by the 24mer after two cycles of low stringency annealing. If the reverse transcription primer is effectively removed from the PCR then the background of cDNAs which are not differential may be diminished.

(b) Autoradiograph of differential display fingerprints of 9.5 *dpc* embryo RNA using the dT<sub>24</sub> primer for reverse transcription and PCR with the protocol outlined in figure 4.1a. The reaction products were separated by electrophoresis through a 5% (w/v) native polyacrylamide gel as before. The tracks represent the products seen when the high stringency annealing cycles are carried out at 45°C, 50°C, 55°C and 60°C respectively. As the annealing temperature rises above 50°C, reaction products disappear indicating that the dT<sub>24</sub> primer does not participate efficiently in the reaction.

(a)



(b)

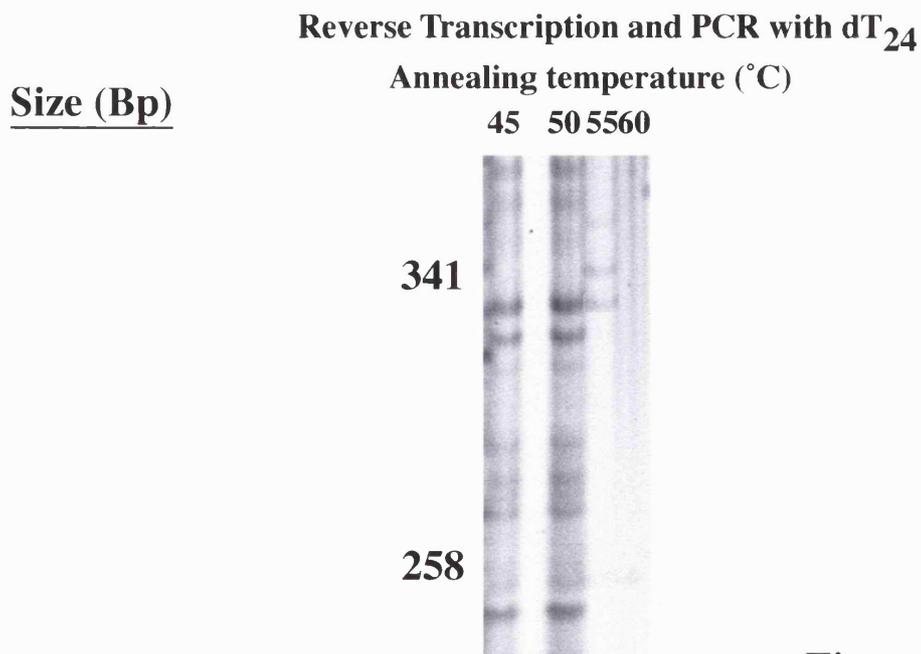


Figure 4.1

strand cDNA, the annealing stringency can be raised so that only PCR products with the arbitrary primers at their ends are amplified and those with the reverse transcription primer can be excluded. This should remove the problems of multiple clones arising from single bands and of clonable species being amplified from apparently blank areas of the differential display gel if these problems are indeed due to the participation of the reverse transcription primer in the PCR step. The idea of using longer primers in the differential display has been postulated before (Ayala *et al.*, 1995; Diatchenko *et al.*, 1996a), though it does not appear to have been widely adopted. AP-PCR approaches use long arbitrary primers to generate differential PCR amplicons, but they involve reverse transcription with one of the primers used, and so the same theoretical problems arise. The use of longer primers in the display PCR should also lead to larger reaction products (because there should be fewer sites in a given cDNA where an arbitrary 24mer can anneal), and the reamplification step can be carried out at high stringency, thereby eliminating internal priming. In the protocol proposed here, the RNA population is not subdivided at the time of reverse transcription with anchored primers. This may not be necessary as a reduced number of bands per reaction are predicted by the use of longer primers (Diatchenko *et al.*, 1996a).

In order to ascertain whether it is possible to eliminate bands solely due to the reverse transcription primer from the resulting cDNA fingerprint, 50ng 9.5 *dpc* mouse embryo RNA was reverse transcribed with a dT<sub>24</sub> primer and the products PCR amplified with the dT<sub>24</sub> as the only primer. The reaction conditions included two cycles of low stringency annealing (40°C) followed by 40 cycles of annealing at 45°C, 50°C, 55°C or 60°C. The reaction products were resolved by electrophoresis and an autoradiograph is shown in figure 4.1b. It can be easily appreciated that visible bands disappear at higher temperatures, suggesting that at 60°C the reverse transcription primer has reduced participation in the PCR reaction.

## **5.2. A comparison of mouse somites and presomitic mesoderm**

A test of the modified protocol was set up for the comparison of dissected first born somites and presomitic mesoderm. Larger clones with less background were expected; the experience with differential display analysis of plasmid libraries suggested that each band would represent a differential cDNA and the use of longer primers to generate larger products may allow their direct use for *in situ* hybridisation studies. The steps in the modified protocol are itemised below;

1. Reverse transcription was with a with a dT<sub>24</sub> primer.

2. One eighth of the reverse transcription reaction was used per PCR.
3. PCR was performed after the addition of a single arbitrarily chosen 24mer primer.
4. The PCR annealing temperature was 40°C for two cycles followed by 60°C for 40 cycles.

### **5.2.1 Tissue isolation and RNA preparation**

All dissections were performed by Dr. Dennis Summerbell (Division of Eukaryotic Molecular Genetics, NIMR) on 8 - 12 somite stage mouse embryos in ice cold PBS over a 90 minute period. No enzymatic treatments were used. Only one side of each embryo was dissected. Presomitic mesoderm and the first born somite (defined as being the first somite with a visible caudal margin) were removed. Approximately 25 of each structure were lysed in 100µl denaturing solution and poly (A)<sup>+</sup> RNA was prepared by the standard protocol.

### **5.2.2 RNA quality analysis and differential display**

The equivalent poly (A)<sup>+</sup> RNA from two presomitic mesoderm blocks or ten first born somites (which were assessed as being of similar volume) was reverse transcribed with the dT<sub>24</sub> primer in an otherwise conventional 20µl differential display reverse transcription reaction.

As a rapid quality analysis of the RNA, 2 µl of each reverse transcription was used in a PCR with primers for *oct -3*, *myf-5* and the novel murine *Delta* homologue *Delta S*. (Primers a kind gift of Dr. S. Dunwoodie; this gene is expressed by *in situ* hybridisation in the presomitic mesoderm only up to the 15 - 20 somite stage, after which it appears in the first three somites (un-published data)). The primers for *Myf -5* anneal on either side of the introns in the gene and give a product of 224bp from cDNA and 1547bp from genomic DNA. As seen in figure 4.2a, *oct -3* and *myf-5* were detectable in both samples but *Delta S* was only seen in the presomitic mesoderm. These findings are as predicted for *oct -3* and *Delta S* by *in situ* hybridisation pattern but *myf-5* is only detectable by *in situ* in the somites. However, others have claimed to detect *myf-5* in presomitic mesoderm by reverse transcription PCR analysis (Kopan *et al.*, 1994).

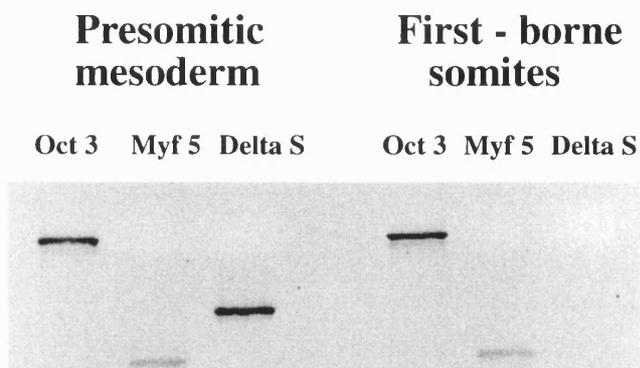
A differential display analysis using the modified protocol with the same cDNAs was performed with the three arbitrary primers BamO4, BamO5 and BamO6. An autoradiograph of the electrophoresed reaction products is shown in figure 4.2b. The

#### **Figure 4.2**

(a) Photograph of the reaction products of a qualitative reverse transcription analysis of RNA from pooled presomitic mesoderm and first born somites from 8 - 12 somite stage mouse embryos after separation by agarose gel electrophoresis. Poly (A)<sup>+</sup> enriched RNA from the equivalent of two presomitic mesoderm tissue blocks and ten first born somites was reverse transcribed with the dT<sub>24</sub> primer in a standard 20µl reaction and 1µl of the resulting cDNA was used as a template for a 40 cycle PCR with three primer pairs specific for the genes *oct-3*, *myf-5* and *Delta S*. The primer pairs were: *oct-3*; OC32 and OC33 (which yield a 562bp product from cDNA only): *Myf 5*; IN1F and IN1R (which yield a 225bp product from cDNA and a 1377bp product from genomic DNA): *Delta S* primers which were a kind gift of Dr. Sally Dunwoodie, give a 342bp product from cDNA and genomic DNA. All three products are detectable in presomitic mesoderm RNA but *Delta S* is not detectable in the somite RNA.

(b) Autoradiograph of a differential display analysis (performed with the modified protocol using arbitrary 24mer primers) of the first born somite RNA and presomitic mesoderm RNA studied in figure 4.2a. The products were separated by electrophoresis through a native polyacrylamide gel run at 50W for two hours at 4°C. Each pair of tracks was generated using the arbitrary 24mer primer indicated. The presomitic mesoderm tracks are shown on the left and the somite tracks are on the right. The clones generated from bands P4 and S6 gave *in situ* hybridisation patterns consistent with their appearance in the gel.

(a)



(b)

Reverse Transcription Primer dT<sub>24</sub>

Size (bp)

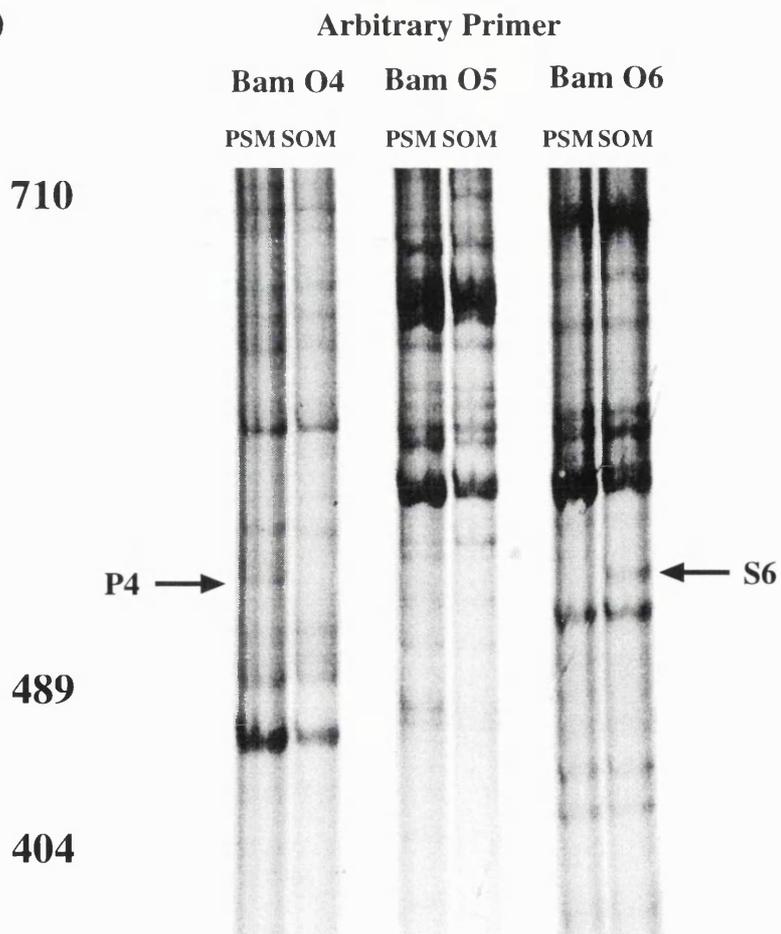


Figure 4.2

overall number of bands was similar to that seen in the first experiments performed on whole embryo RNA, however there were many fewer differential bands. The smallest products observed were in the size range 250 - 300bp with most being in the range 500 - 1000bp. A total of 12 bands (five differential and seven non - differential) were reamplified with the corresponding arbitrary primer only at high stringency for 35 cycles. The reamplification reactions generally gave single products of the expected sizes, the largest being 1646bp. After cloning each band gave rise to a single cloned species on sequence analysis (four to six clones analysed for each band) and, as expected, each clone had the reamplification primer sequence at its insert ends. No evidence of genomic DNA or extraneous plasmid DNA was found.

### **5.2.3 Secondary screening of differential clones**

Five differential bands were seen in this experiment, all of which gave rise to single clones, each corresponding to known murine cDNA sequences. They were tested by whole mount *in situ* hybridisation analysis on a range of post-gastrulation mouse embryos from 4 - 30 somite stages. Only two clones, derived from bands P4 (presomitic mesoderm specific) and S6 (somite specific) gave hybridisation signal.

### **5.2.4 Clone S6**

This 525bp fragment corresponded to a recently cloned gene, Neuronal Leucine Rich Repeat protein 1 (NLRR-1) (Taguchi *et al.*, 1996). It was isolated as a full length species from a neonatal mouse brain cDNA library after a low stringency screen with a human partial cDNA probe generated by random sequencing of ESTs from a human infant brain library (Adams *et al.*, 1993). This probe contained a sequence encoding a leucine rich repeat (LRR) motif. This is a protein domain consisting of predominantly leucine or other aliphatic amino acid domains that exist as iterated arrays in a wide range of usually membrane associated gene products present in virtually all eukaryotes (reviewed in Kobe and Deisenhofer, 1994). The aim of the workers concerned was to isolate LRR proteins that may have roles in brain development.

NLRR-1 was isolated as a 3.7Kb full length clone containing an entire open reading frame. It has been shown to contain a LRR domain similar to those of the *Drosophila* genes *slit* and *tartan*, with cysteine rich amino - and carboxyl - flanking regions on either side of the LRR. Unlike *slit* however, NLRR-1 has no EGF repeats (Rothberg *et al.*, 1988; 1990). Hydrophobicity studies of NLRR-1 have shown it to carry a signal peptide and a transmembrane domain at its carboxyl end. The workers who cloned NLRR-1 found it to be clearly expressed in 15.5 *dpc* embryos on parasagittal

section in the central nervous system (diffusely in the brain and spinal cord and in dorsal root ganglia). In the adult brain by similar radioactive *in situ* hybridisation it was seen in the cerebellum and dentate gyrus of the hippocampus. Evidence by Northern blot of expression at 11.5 *dpc* was shown but no early *in situ* hybridisation data were provided.

S6 contained entirely coding sequence corresponding to positions 2067 - 2592 of NLRR-1 encompassing the region between the carboxyl terminal flank of the LRR and the transmembrane domain. The *in situ* hybridisation data for this probe are shown in figure 4.3a. Expression is first seen faintly in 4 somite stage embryos in the dorsal region of all somites but not the presomitic mesoderm. By 8 somites expression is more clearly established again in the dorsal somite but also in the forebrain. At the 12 somite stage (from which the original dissections were made) this pattern is maintained and persists right through to the 35 somite stage. Close analysis of 25 somite stage embryos shows expression in the first born somite but not the presomitic mesoderm. No other central nervous system expression is detected other than in the forebrain as before. This clone provides a clear example of a gene isolated by the modified differential display protocol the expression of which follows its expected pattern from the display gel.

#### **5.2.5 Clone P4**

This presomitic mesoderm specific clone carried a 514bp insert corresponding to the gene for the transmembrane channel protein, connexin 43 (Sullivan *et al.*, 1993). Connexin 43 has been well studied in early mouse embryos by radioactive section *in situ* hybridisation analysis and in zebrafish embryos by whole mount *in situ* hybridisation (Essner *et al.*, 1996). It is known to be expressed in early mesoderm and ectoderm in primitive streak stage embryos and early post-gastrulation stages but by 16 - 20 somite stages, it appears in the youngest somites and various central nervous system structures. These data are reproduced in figure 4.3b. In 12 somite embryos corresponding to the stage at which the dissections were made, expression is seen at high levels in the ectoderm overlying the tailbud and at lower levels in the presomitic mesoderm but not in the somites. Somitic expression is apparent in the earliest somites by the 20 somite stage. This is a further example of the origin of the band in the differential display predicting the pattern of expression of the cloned gene by *in situ* hybridisation.

### **Figure 4.3**

*In situ* hybridisation analysis of clones S6 and P4.

(a) Clone S6. Digital images of embryos between the 4 and 35 somite stages. Expression is first seen in the somites of 4 somite stage embryos (s). It is markedly restricted to the dorsal somite. By the 12 somite stage, midline expression is seen in the forebrain (f). No expression outside of the somites or forebrain is seen up to the 35 somite stage. Detailed analysis of 25 somite stage embryos shows expression in the first borne somite (1st). The triangular domain of expression in the mature somites at this stage is similar to that of *Myf-5*.

(b) Clone P4. Digital images of embryos between the 12 and 30 somite stages. Early expression is seen in the forebrain (f) and the primitive ectoderm (e) and mesoderm (m) of the tailbud. By the 20 somite stage, expression is visible in the most caudal 8 - 10 somites (ls).

**(a)**



**4, 8, 12 and 16 somite stages      20 somite stage**

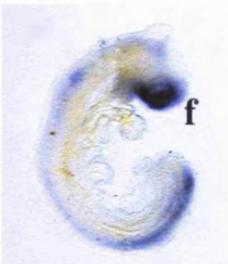


**35 somite stage**



**25 somite stage**

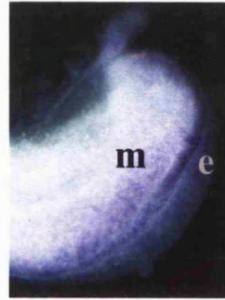
**(b)**



**12 somite stage**



**16 somite stage**



**20 somite stage**



**30 somite stage**



**Figure 4.3**

### **5.3. Refinements in template preparation**

The results from this initial somite/presomitic mesoderm experiment were encouraging; the bands now gave single clones and these were large enough for direct secondary screening by *in situ* hybridisation analysis. Only two out of five bands tested gave predicted hybridisation patterns but no false positive or ubiquitous clones were isolated. The finding of a majority of clones giving no *in situ* hybridisation pattern was in keeping with the earlier study on the plasmid libraries. In order to ascertain the feasibility of carrying out this analysis on even smaller dissected structures, it was decided to use differential display to compare dissected individual somites and presomitic mesoderm from single embryos. It is possible that such an approach may lead to even greater sensitivity in the detection of differential clones.

Dissections were carried out as before to isolate individual presomitic mesoderm, first born somite, and somites 2 -4 and 5 - 7 *en bloc* from 10 somite stage embryos as shown in figure 4.4. Two identical 10 somite stage embryos from the same litter were chosen; from one somites 5 - 7 and the contralateral presomitic mesoderm were removed; from the other, ipsilateral first born somite and somites 2 - 4 were taken. In addition, somites 2 - 4 and 5 - 7 were dissected from two identical 20 somite stage embryos. Each structure was lysed in 40µl denaturing solution and poly (A)<sup>+</sup> RNA was prepared by the standard protocol with a final elution volume of 20µl.

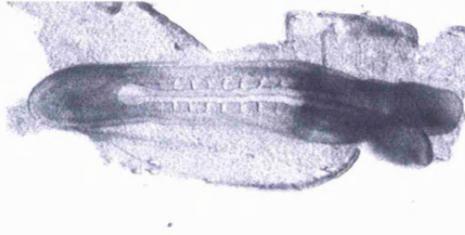
#### **5.3.1 Differential display of individually dissected embryonic structures**

2µl presomitic mesoderm RNA or 10µl somite RNA were reverse transcribed as before in a 20µl volume with the dT<sub>24</sub> primer and 1µl of the reaction used for PCR analysis with primers detecting *HPRT*, *myf -5*, *oct -3*, *myogenin* and *DeltaS*. The PCR profiles of the respective cDNAs are shown in figure 4.4 below each embryonic structure. It should be stressed that these analyses are in no way meant to be quantitative and act as a rapid RNA quality check. 45 cycles of PCR were performed and the products visualised after agarose gel electrophoresis. Relative differences in product abundance within a given tissue could be due to the relative efficiencies of the primers for PCR (as is the case for *myf -5*, the primers for which are poor). As expected, *HPRT* and *oct -3* were detectable in all tissues as was *myf -5*, in keeping with the findings of others (Kopan *et al.*, 1994). Reassuringly, *myogenin* was only detectable in cDNA from somites 5 - 7 in both the 10 and 20 somite stage embryos and not in less mature somites or presomitic mesoderm, in keeping with its *in situ* pattern. The detection of *Delta S* is also in keeping with its known expression pattern,

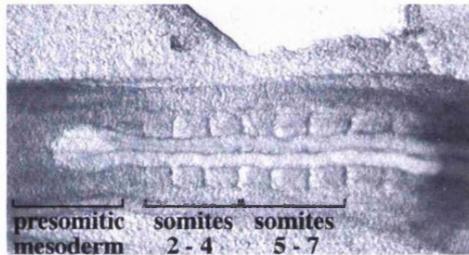
#### **Figure 4.4**

Digital images of and photographs of the reverse transcription PCR fingerprints derived from a 10 somite stage mouse embryo and dissected presomitic mesoderm and somites 2 - 4 and 5 - 7 from 10 - and 20 - somite stage embryos. Poly (A)<sup>+</sup> enriched RNA was purified from a single whole embryo in a volume of 40µl and one of each of the dissected embryonic structures in a volume of 20µl. 1ml whole embryo RNA, 2µl presomitic mesoderm RNA or 10ml somite RNA were reverse transcribed in a 20µl volume with the dT<sub>24</sub> primer, and 1ml of this reaction used for PCR analysis with primers specific for *HPRT*, *myf -5*, *oct -3 myogenin* and *DeltaS*. The details of these primers have been given before.

Whole 10 somite stage embryo



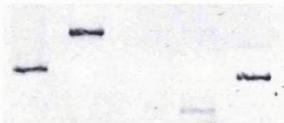
HPRT Oct-3 MGN Myf-5 Delta-S



Presomitic mesoderm



HPRT Oct-3 MGN Myf-5 Delta-S

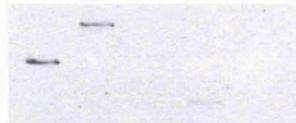


10 somite stage

Somites 2 - 4

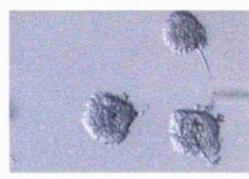


HPRT Oct-3 MGN Myf-5 Delta-S

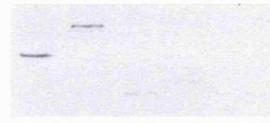


10 somite stage

Somites 5 - 7

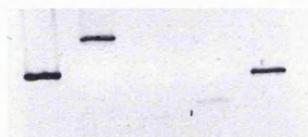


HPRT Oct-3 MGN Myf-5 Delta-S



10 somite stage

HPRT Oct-3 MGN Myf-5 Delta-S



20 somite stage

HPRT Oct-3 MGN Myf-5 Delta-S



20 somite stage

**Figure 4.4**

being seen in the presomitic mesoderm only in the 10 somite stage mice but also appearing in the earliest somites by the 20 somite stage.

An autoradiograph of the differential display gel obtained when these cDNAs were amplified with BamO5 is shown in figure 4.5. The fingerprints obtained were almost identical with very few real differences. Only bands that were clearly differential were studied further as it was not clear whether the rough estimates that had been made of the RNA that could be extracted from the tissues were accurate enough to allow the system to detect up - or down - regulated cDNAs after 40 cycles of PCR. Possibly on account of the very small amount of material being used here, some bands that appeared superficially differential could be seen on close inspection in adjacent tracks (examples of this are highlighted by asterisks in figure 4.5). Only two bands were unambiguously differential and both were somitic (marked as S53 and S54). S54, which was cloned from a band specific to somites 5 - 7 of both 10- and 20-somite stage embryos, encoded a novel sequence with no homologies at the DNA level . It failed to give an *in situ* hybridisation signal.

### **5.3.2 Clone S53**

S53 was specific to somites 5 -7 of 20 somite stage embryos. It encoded the murine homologue of a gene known as NSP-C (neuroendocrine specific protein) in humans or RESP18 (regulated endocrine specific protein) in rats. RESP18 was cloned from a rat *pars intermedia* pituitary cDNA library on account of its parallel regulation with proopiomelanocortin (Schiller *et al.*, 1995), and NSP-C after screening of a human small cell lung cancer cell line expression cDNA library with antibodies raised against various tumour markers (Senden *et al.*, 1996). The 21KD gene product has no homology to any known proteins and no obvious function. It has been identified as a neuroendocrine ER (endoplasmic reticulum) protein and appears to be localised to the ER lumen by a totally novel mechanism. It is not secreted except in certain pituitary tumour cell lines, and appears to be proteolytically degraded in a post - ER, pre - Golgi compartment (Schiller *et al.*, 1995). No embryonic studies of its expression have been reported.

*In situ* hybridisation expression of S53 over the 5 - to the 25 - somite stages is shown in figure 4.6. It is first seen in the earliest stage in the neural crest and its migrating derivatives. By the 20 somite stage expression is clearly visible in the older somites (from about the fifth somite) confirming the prediction of its differential display gel appearance. This expression probably represents neural crest cells as they migrate through the rostral sclerotome of the somites, though section data are still awaited for

#### **Figure 4.5**

Autoradiograph of a differential display analysis of RNA derived from the presomitic mesoderm, first borne somite, somites 2 - 4 and somites 5 - 7 of a 10 somite stage embryo and somites 2 - 4 and 5 - 7 from a 20 somite stage embryo. The modified protocol was used with reverse transcription using dT<sub>24</sub> and PCR with BamO5. The reaction products were sparated by electrophoresis through a 5% (w/v) native polyacrylamide gel. The bands marked by asterisks, although apparently differential, were not so, and on close inspection could be seen to have sister bands in adjacent tracks. The bands S53, which was generated from somites 5 - 7 of 20 somite stage embryos only, and S54, which was generated from somites 5 - 7 of 10 - and 20 - somite stage embryos were selected for further study (see text).

Reverse Transcription Primer dT<sub>24</sub>  
Arbitrary Primer Bam 05

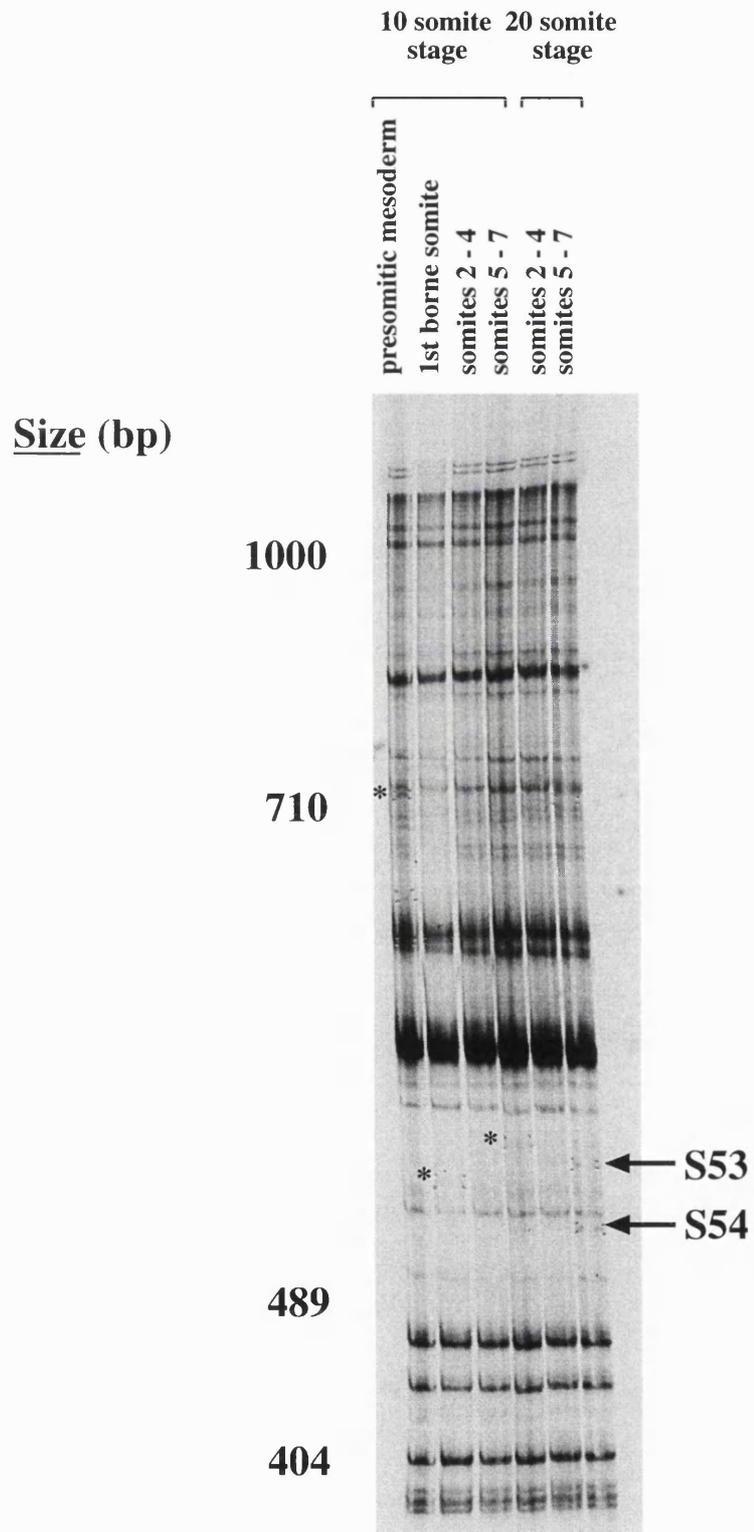
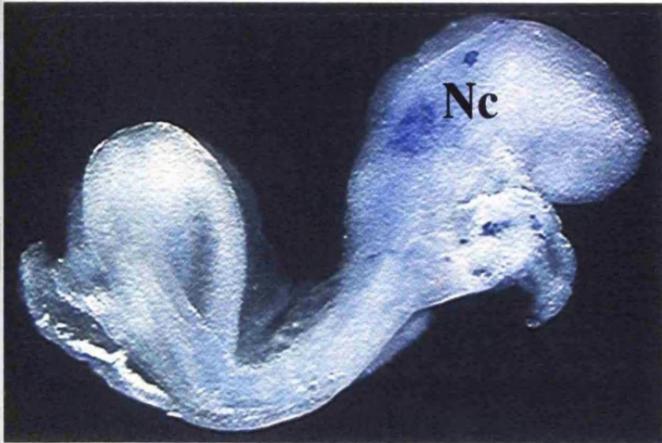


Figure 4.5

**Figure 4.6**

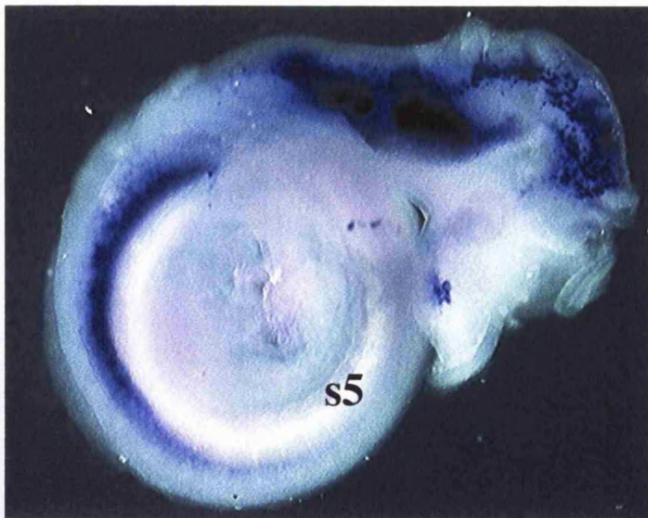
Digital images of an *in situ* hybridisation analysis of the expression of S53 between the 5 - and 30 - somite stages. Initial expression at the 5 - somite stage is in the neural crest (Nc) and subsequently in its derivatives. By the 20 - somite stage, expression is first visible from about the fifth somite (s5). By the 25 - 30 somite stage, striking expression is seen in the dorsal root ganglia (DRG). From the 20 - somite stage onwards, foci of expression are seen in what are probably cranial nerve nuclei and nerve tracts (Nt).



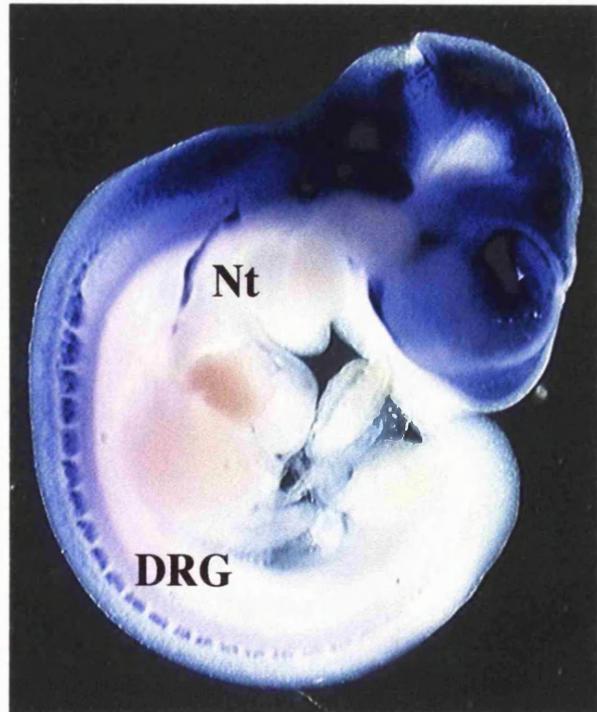
**5 somite stage**



**12 somite stage**



**20 somite stage**



**25 - 30 somite stage**

**Figure 4.6**

this clone. By 30 - 35 somites this “somitic” expression is striking but appears to be localised to the dorsal root ganglia which lie adjacent to the somites. It would have been impossible to separate the neural crest derivatives from the somite proper at the time of dissection without enzymatic digestion. This clone was isolated from a differential band generated from the RNA of somites 5 - 7 of 20 somite embryos and it encoded a gene whose first somitic expression was seen at this stage in these somites by *in situ*.

#### **5.4 Isolation of homologues of NLRR-1**

The clone S6 (NLRR-1) gave an intriguing pattern of hybridisation in the somites highly reminiscent of the bHLH transcription factor *myf-5* whose transcriptional regulation has been a long standing interest of the laboratory. Preliminary section studies of the S6 wholemount embryos not presented here suggest that it is restricted to the dorsal myotome. As this class of gene exists in closely related families in a wide variety of species (Rothberg *et al.*, 1990), it was decided to isolate other homologues of this type of LRR gene from embryonic RNA.

##### **5.4.1 A 5'RACE strategy**

The workers who isolated NLRR-1 also published the sequence of a related LRR containing gene isolated in the same way which they named NLRR-2 (Taguchi *et al.*, 1996). This is a 217bp EST sequence which is expressed in the pyramidal layer of the hippocampus of the adult mouse brain, but for which no convincing embryonic expression has been shown. They subsequently published a third related gene, NLRR-3, isolated again in a similar fashion from adult mouse brain which is upregulated in damaged adult murine cerebral cortex.

Figure 4.7a shows the protein alignments of NLRR-1 and NLRR-3 along with the conceptual translation of the EST NLRR-2 in the region in which they overlap *i.e.* the LRR domain. From this analysis, nested degenerate oligonucleotide primers corresponding to the protein sequences KELG(I/L)NNM (primer KELG; 21-mer; 64 - fold degeneracy) and GDF(K/A/S)NML (primer GDFX; 20-mer; 128 - fold degeneracy) were synthesised and used in a 5' Rapid Amplification of cDNA Ends (RACE) experiment to isolate novel related cDNAs from somitic RNA (figure 4.7a).

5µl RNA from somites 2 - 4 and somites 5 -7 of 20 somite stage embryos (as shown in figure 4.4) were reverse transcribed with 835ng KELG primer (corresponding to 13ng each individual oligonucleotide) in a 20µl reaction volume under standard

#### **Figure 4.7**

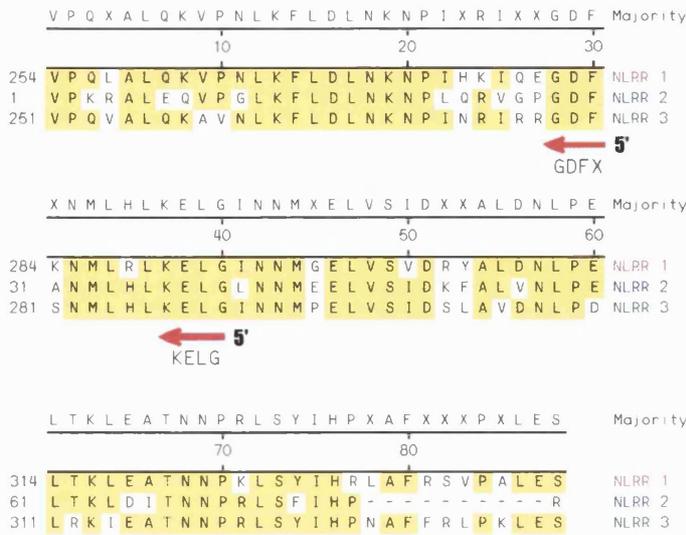
(a) Protein alignment of the genes NLRR-1, NLRR-3 and the 217bp EST NLRR-2 over the region of the LRR. Conserved residues are highlighted in yellow. The positions of the primers KELG and GDFX are marked with arrows which show their orientation relative to the coding sequence.

(b) DNA sequence of the LRR-4 clone. The underlined sequence represents a continuous open reading frame; the corresponding protein sequence is given underneath. The position of the GDFX primer at its 3' end is highlighted in red and marked with the arrow and the sequence highlighted in blue shows a region of 93% identity with NLRR-2.

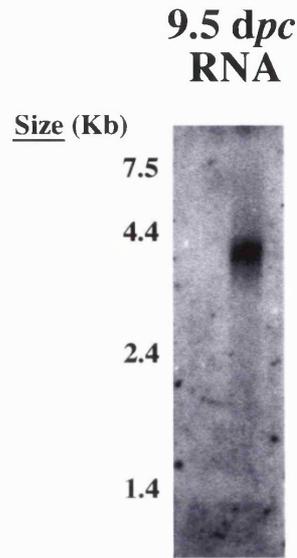
(c) Phosphorimage of a northern blot containing 5µg poly (A)<sup>+</sup> enriched RNA from 9.5 dpc mouse embryos after hybridisation with an antisense probe for LRR-4. A strong signal is seen at approximately 4Kb.

(d) Protein alignment of the genes NLRR-1, NLRR-3 and the LRR-4 over the region of the LRR and part of the amino flanking region. Conserved residues are highlighted in yellow.

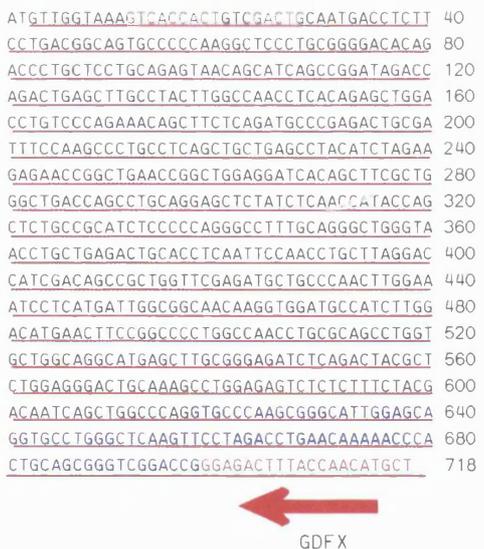
(a)



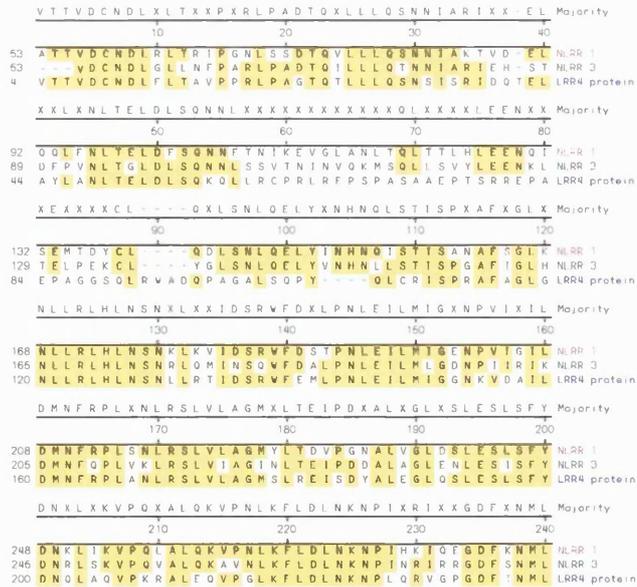
(c)



(b)



(d)



VGKVTTVDCNDLFLTAVPPRLPAGTQTL LLSNSISRIDQ 40  
 TELAYLANL TELDLSQKQLLRCPRLRFPSPASAAEPTSRR 80  
 EPAEPAGGSQLRWADQPAGALSOPYQLCRISPRAFAGLGN 120  
 LLRLHLNSNLLRTIDSRWFEMLPNLEILMIGNKVDAIDL 160  
 MNFRPLANLRSVLGMSLREISDYALEGLQSL SLSFYD 200  
 NQLAQVPKRALEQVPG LKFLDLNKNPLQRVGP GDFTNM 238

Figure 4.7

conditions for one hour. The reaction was made up to 40µl with 2x terminal transferase buffer and a dA tailing reaction performed for 15 minutes. After extraction with phenol and chloroform followed by ethanol precipitation, the purified tailed cDNA was used as a template for PCR amplification with the primers GDFX and EcodT using the parameters 94°C for one minute; 2 cycles of 94°C for 30 seconds, 40°C for two minutes and 73°C for one minute; 35 cycles of 94°C for 30 seconds, 53°C for one minute, 73°C for 45 seconds.

Reaction products were seen only from the cDNA of somites 5 - 7. They were end polished and size fractionated to a range of 500 - 1000bp after agarose gel electrophoresis before cloning. After *lac* screening, 36 recombinant clones were sequenced; 28 corresponded to the 5' end of NLRR-1 and 8 encoded a novel 718bp fragment, named LRR-4 with a single open reading frame that represented a related gene. The DNA sequence of this species and its conceptual translation is shown in figure 4.7 b. At the 3' end the sequence of the GDFX primer is highlighted as is a region of 101bp which has 93% identity at the DNA level with NLRR-2.

#### **5.4.2 Analysis of LRR-4**

By Northern blot analysis, a LRR-4 riboprobe gave a strong signal at approximately 4Kb in 9.5 *dpc* RNA (figure 4.7c). Protein sequence alignments of LRR-4 with NLRR-1 and NLRR-3 (figure 4.7d) showed it to have 60.5% similarity with NLRR-1 and 52.4% similarity with NLRR-3 over the LRR and carboxyl end of the amino flanking region. The conserved periodicity of the hydrophobic residues in NLRR-1, LRR-4 and *Drosophila slit* is shown in figure 4.8a, and the conserved residues of the amino flanking region of these genes and NLRR-3 are shown in figure 4.8b. The sequence of the LRR-4 fragment is truncated in this latter domain.

Preliminary wholemount *in situ* hybridisation studies have been performed with LRR-4 (see figure 4.9). The earliest embryo studied here was at the 15 - 20 somite stage (figure 4.9a), where expression was seen in mature somites. This was maintained through development to the 30 somite stage (figure 4.9b, c). At the 20 - 25 somite stage, the somitic expression was classically myotomal, though section data are not yet available. Diffuse central nervous system expression was observed with this probe and a clear boundary of expression was seen in the midbrain (figure 4.9d).

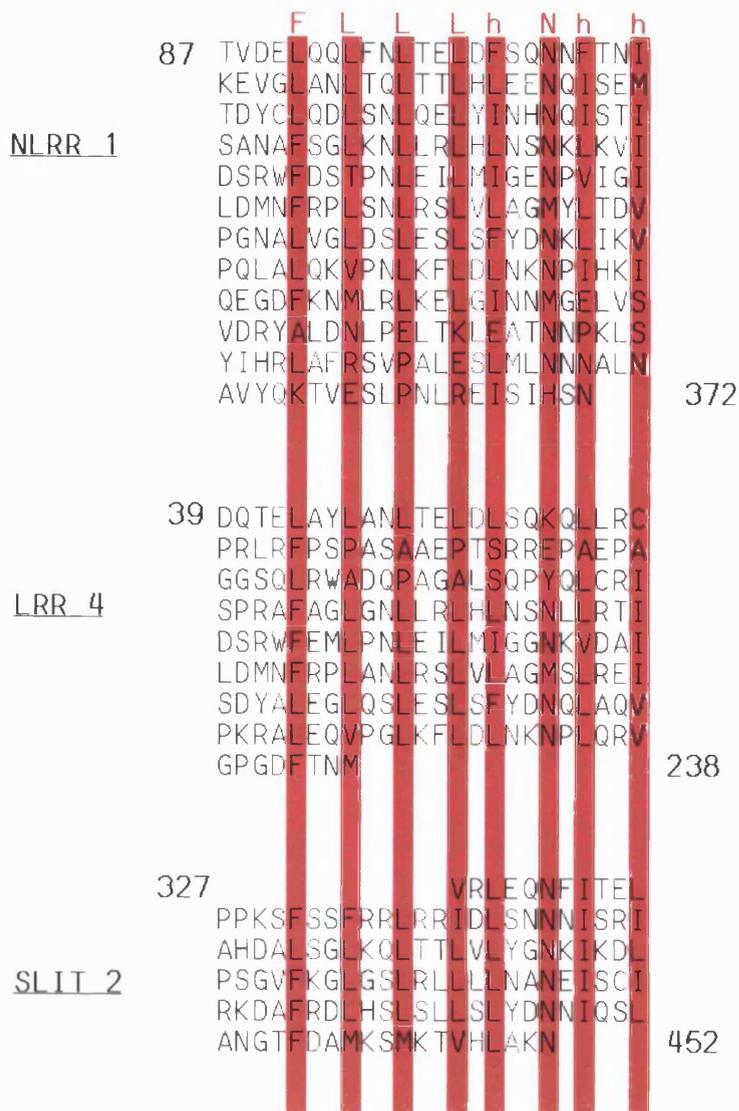
**Figure 4.8**

(a) Alignment of the conserved residues in the LRRs of NLRR-1, LRR-4 and the second LRR of the *Drosophila Slit* gene (*slit 2*). Consensus residues are shaded and the commonest amino acids at each position of the consensus are given in red (F phenylalanine, L leucine, N asparagine, h hydrophobic).

(b) Alignment of the amino flanking regions of NLRR-1, NLRR-3, the second LRR of the *Drosophila Slit* gene(*slit 2*) and the partial amino flanking region of LRR-4. The majority sequence is given at the top with highly conserved residues highlighted in red and boxed in the alignment. Conserved residues are highlighted in the alignment in yellow.

## Leucine rich repeat regions

(a)



(b)

## Amino flanking regions

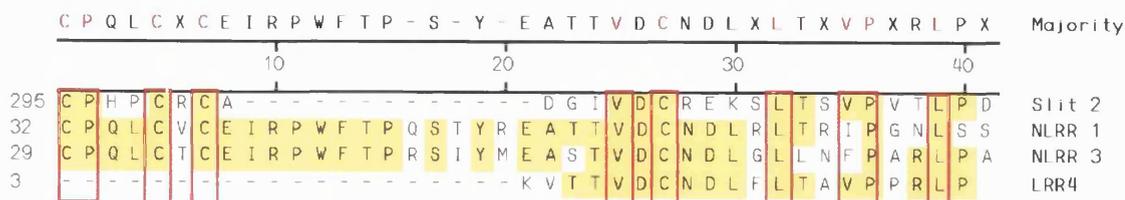


Figure 4.8

**Figure 4.9**

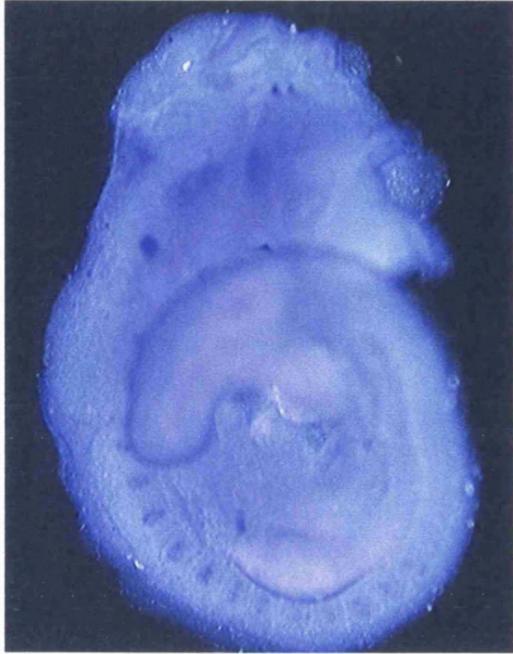
Digital images of the *in situ* analysis of the expression of LRR-4 in embryos between the 15 - and 30 - somite stages.

(a) At the 15 - 20 somite stage, expression is seen from about the seventh somite in a classically myotomal distribution.

(b), (c) 20 and 20 - 35 somite stages, clear expression is seen in the same domain from about the third somite

(d) Detailed analysis of the somitic expression in 20 - 25 somite stage embryos revealing a myotomal pattern. A clear boundary of expression (b) in the midbrain is seen at this stage.

**(a)**



**15 - 20 somite stage**

**(b)**



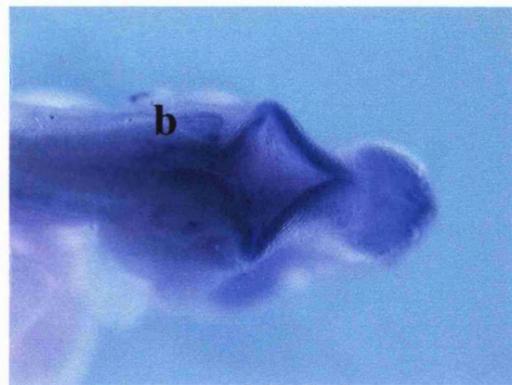
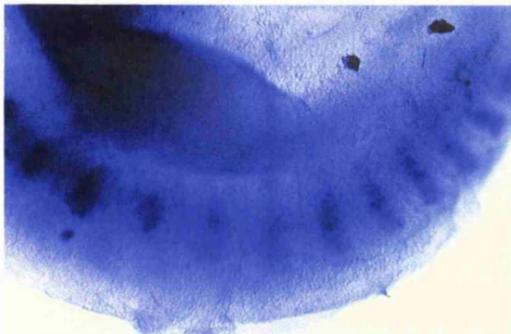
**20 - 25 somite stage**

**(c)**



**25 - 30 somite stage**

**(d)**



**20 - 25 somite stage**

**Figure 4.9**

## 5.5 Summary

These studies showed that it is technically feasible to perform a differential display analysis of dissected murine embryonic structures using *in situ* hybridisation as a secondary screen. The problem of multiple clones arising from a single band was overcome by preventing the participation of the reverse transcription primer in the PCR. This was achieved by using a single long (24mer) arbitrary primer which permitted an increase in the annealing temperature of the PCR after two initial low stringency cycles. The use of longer primers ensured that longer clones of a size suitable for *in situ* hybridisation studies could be obtained; not only did they allow the generation of longer bands *per se* but they reduced the problem of internal priming during the reamplification step as this could be carried out at high stringency throughout.

## CHAPTER 6

### Discussion

Since its inception in 1992, differential display has found applications in all areas of molecular biology as a tool for rapidly isolating differentially expressed genes. At the recent Cold Spring Harbour Symposium on differential display and related technologies, it was commented that at that time, over a period of four years, some 370 publications had been made in the scientific literature reporting the successful application of the technique or modifications designed to improve it.

The method involves the generation of cDNA fingerprints from the RNAs of two (or more) tissues under study after reverse transcription and PCR under low stringency with short arbitrary primers. Each DNA band in the fingerprints is proposed to represent a given gene expressed in the starting sample and a side by side comparison will lead to the identification of species specific to one tissue. After cloning, the cDNAs can be tested in some secondary screen to prove whether or not they are truly differentially expressed.

For a mouse embryologist, who has to deal with minute quantities of material and consequently for whom RNA availability is limiting, the technique offers a tantalising route to the isolation of novel genes from dissected embryonic structures without any *a priori* information regarding their sequence. It is rapid, can operate with very small amounts of RNA and the fingerprints obtained are generally reproducible for a given RNA preparation. In addition, many authors have suggested that the method tends to identify low abundance species (Conway 1995; Guimarães *et al.*, 1995a,b; Mathieu-Daudé *et al.*, 1996b; Rafaeloff, 1996). The secondary screen that is generally employed to show the veracity of a differential band has classically been that of Northern blotting, however for the mouse embryologist this is generally not possible. Instead the highly informative and aesthetically pleasing method of *in situ* hybridisation has developed as a rapid way of localising the site(s) of expression of a given transcript in an embryo.

Despite its potential, there have as yet been no published attempts to use differential display in the mouse embryo that have either compared dissected embryonic structures or that have used *in situ* hybridisation as a secondary screen. The reasons for this are manifold, but the inherent technical problems that plague the method as it was originally devised, play a large role in this.

The work presented here has attempted ultimately to use differential display in a comparison of gene expression between dissected murine embryonic structures (presomitic mesoderm and the first - born somite). No attempt has been made to address the issue of how many expressed genes from the total cDNA population of a tissue one could display, and its use merely as a method of screening for differential gene expression has been studied. The problems inherent in the original protocol have been exposed and steps to circumvent them developed.

## **6.1 The generation of reverse transcription PCR fingerprints**

The original protocol of Liang and Pardee, (1992) was initially used here to generate differential display products. Two minor modifications were made at the outset and retained throughout. In keeping with NIMR safety recommendations,  $^{33}\text{P}$  was used as the radio isotope of choice.  $^{35}\text{S}$  is believed to lead to the generation of volatile radioactive thiols during PCR (Bauer *et al.*, 1993; Liang *et al.*, 1995). Secondly, native polyacrylamide gels were used to display reaction products; it is well established that native gels are adequate in the resolution of PCR products and when compared to denaturing gels, give less complex footprints with comparable numbers of differences (Bauer *et al.*, 1993; Liang *et al.*, 1995). Furthermore native gels are easier to manage after electrophoresis and eluted DNA is readily reamplified without the need for ethanol precipitation.

### **6.1.1 Template preparation**

Commonly used RNA extraction protocols vary, the particular biological system under study often dictates which methods can or cannot be used. The acid - guanidinium - phenol method (Chomczynski and Saatchi, 1987) is a favoured technique for embryologists. It is rapid and can be easily scaled down to situations where only a few cells are available (Brady and Iscove, 1993; Dulac and Axel, 1995; Harrison *et al.*, 1995). It has one major drawback: the RNA is separated from DNA, proteins and other cellular debris by pipetting after a phase separation step carried out in a microcentrifuge tube. The aqueous phase containing RNA has genomic DNA at its interface with the organic phase. As a consequence, there is always a risk of DNA contamination of the RNA prepared and as the extraction volumes diminish, the risk increases. In a PCR based differential screening procedure, this is a problem because, unless genomic DNA contamination is exactly equal amongst all the samples under study, then it is predicted that at least some differential bands would be due to contaminants. Figure 2.1 illustrates this point. Fingerprints are clearly seen when non - reverse transcribed total RNA prepared by the acid - guanidinium - phenol method is subjected to a low

stringency PCR with arbitrary primers, and these are largely curtailed after poly (A)<sup>+</sup> enrichment.

Figure 2.2 shows that poly (A)<sup>+</sup> enrichment removes some species from the fingerprint of total RNA which are only visible after reverse transcription. Presumably these represent cDNA species from non - polyadenylated RNAs *e.g.* ribosomal RNA. Such species were never cloned and sequenced and so their identities are not known for certain. However, sequence analysis of clones from bands that were made from poly (A)<sup>+</sup> enriched RNA has so far failed to reveal sequence that is known to be anything other than that derived from mRNA.

### **6.1.2 The reproducibility of display patterns**

It is an important feature for the credibility of the method that the arbitrarily primed PCR fingerprints which are obtained from any one tissue type are shown to be reproducible, and consequently that the differential bands chosen for further study would be repeatedly generated as differential if the assay were to be performed on several occasions. The findings here were that the fingerprints for any single poly (A)<sup>+</sup> enriched RNA preparation were reproducible for any given primer pair at a given annealing temperature and magnesium concentration using the thermal cycler and PCR tubes available in the laboratory. Preparations from the pooled embryonic tissues of the same chronological stage made on different occasions gave similar but not identical patterns, probably reflecting the heterogeneity of embryos at the same temporal stage *post coitum* and the varying efficiency of the poly (A)<sup>+</sup> enrichment process.

The conventional approach to achieving this requirement has been to perform the same differential display PCR on a given RNA several times and load the products in adjacent tracks in a gel (Liang and Pardee, 1992). Only bands that are differential in each reaction are chosen for analysis. An alternative approach has been to carry out the PCRs with serial dilutions of input RNA (McClelland *et al.*, 1995). This idea is based upon the observation that unreproducible bands tend to disappear as the amount of starting material is reduced (Zhao *et al.*, 1996; Dr. M.McClelland, personal communication). In the experiments performed here, figure 2.3 shows that reducing the amount of starting material causes some bands to disappear in a given differential display reaction. Others appear to be enhanced. However, the majority of bands remain after 40 cycles of PCR. In general, multiple reactions were not performed for analyses of embryonic RNA as very small amounts of sample were used. This also reduced the amount of radiolabel consumed.

### **6.1.3 The complexity of the display patterns**

The number of bands generated per reaction is a complex function of the primers, the annealing temperature, the magnesium concentration, the thermal cycler and the PCR tubes used. The aim here was not to display the entire cDNA repertoire of each tissue studied, but to investigate whether fingerprint differences observed could be translated into *in situ* hybridisation patterns representing differentially expressed genes. The reaction parameters employed were not varied appreciably except on the occasions where long 24mer arbitrary primers were used in the PCR. Generally an adequate number of bands were seen in each reaction to give a reasonable number of candidates for secondary screening. If a different brand of thermal cycler or PCR tube was used, it is possible that different reaction parameters would be required to generate a manageable number of bands. This in itself should not cast doubt on the validity of the method as long as differential bands could be shown to represent differentially expressed genes.

The numbers of differences observed using the same reaction parameters and primer types varied according to the types of templates analysed. Using the conventional differential display protocol, there were fewer shared bands and more differential bands per track when representative RNAs from the libraries were studied than when pooled 7.5 *dpc* and 8.5 *dpc* embryonic RNAs were examined (compare figures 3.1 and 2.7). There are many possible reasons for this discrepancy. Firstly, the libraries were made from RNAs derived from cleanly dissected anatomical structures in embryos of the same anatomical stage. During the cloning process and subsequent bacterial amplification, differences and similarities between the two may have become amplified or diminished. The embryonic RNAs on the other hand were generated from embryos separated solely by the criterion of the presence or absence of somites. The starting materials were much more heterogeneous, and there may well have been more overlap in gene expression between the two. It is of interest to note that when anatomically distinct mesoderm derived structures (*i.e.* presomitic mesoderm and the first born somite) were compared using the modified protocol, the fingerprint patterns are similar with proportionately fewer differences. This was true whether the RNA was extracted from tissues pooled from several embryos with the same number of somites (figure 4.2) or individual structures from single embryos (figure 4.5).

## **6.2 Analysis of differential bands**

In the work presented here, the reamplification and cloning of differential bands proved initially to be somewhat problematic. When the conventional approach of reamplifying

the bands with the same short primers used in the differential display was employed, the yield of product obtained was low. As the same low stringency conditions were applied to the reamplification reaction in order to ensure adequate primer annealing and product yield, internal priming was observed on larger templates which often yielded smaller reamplification products. In addition, when the bands were cloned, a high percentage of clones was found to represent cloning artefacts. This latter problem was overcome by PCR screening of recombinant colonies and this could be achieved if reamplification was carried out with differential display primers that bore 5' extensions (Liang *et al.*, 1994). The longer primers gave greater yields of product which also facilitated cloning. The problem of internal priming persisted as the first two cycles of the reamplification reaction with the extended primers was also at low stringency.

The prediction of the Liang and Pardee schematic model of differential display (figure 1.1) is that the clones isolated by the method would be directionally cloned and biased towards the 3'UTR with the reverse transcription (dT) primer sequence forming the 3' end of the cDNA amplicon and the arbitrary primer at the 5' end. This rather simplistic view has been shown by many workers to be incorrect. Either primer sequence may reside at each end of the cDNA and only 20 - 30% of clones have the primer sequences in the positions predicted. The low annealing stringencies of the PCR can allow the dT primer to initiate second strand synthesis during the first round of PCR. As has been observed by others (Liang *et al.*, 1992; Guimarães. *et al.*, 1995a) the findings throughout this study have suggested that the 3' most four to six bases of the primers are usually exactly homologous to the target sequences.

### **6.2.1 False clones**

Initial studies of differential display gels generated using the conventional protocol of Liang and Pardee provided evidence that each track contained a background of amplified cDNAs which may or may not have been visible but which could be amplified whenever a *bona fide* band was cut out and reamplified. The first suggestion of this stemmed from the observation that even blank areas of the gels could be "reamplified" and products cloned from them. More worrying was the observation that multiple clones could be generated from even cleanly excised bands including clones of identical size but different sequence. This phenomenon was observed for differential and non - differential bands.

### **6.2.2 Identification of true positive clones**

An inkling that a given differential band might actually contain a species that could be cloned and shown to be differential was provided by data from non - differential bands. Clones generated from 10 non-differential pairs of bands in figure 2.7 showed that multiple sequence types were generated from a given pair of sister bands. However in every case, one clone from each band was identical to one derived from its sister.

The use of representative RNAs derived from plasmid libraries in a differential display analysis allowed a test of the hypothesis that at least one of the clones generated from a differential band might actually be differential. Southern blot filters containing digested plasmid library DNA could be used in an analogous fashion to northern blots and provide a rapid screen for the many clones that could be derived from a single band. Figure 3.2 showed that differential species could be found amongst the clones obtained from a differential band and that the principle of differential display was correct; differential display appeared to be more than just a method for generating random cDNA species. It should be borne in mind that no occurrence of a false positive clone (*i.e.* one that was represented in the *opposite* library to the one predicted by the differential display) was seen.

### **6.2.3 Differential library screening by arbitrarily primed PCR**

The aim of this part of the work was to ascertain whether any of the clones isolated from a differential band were actually differential using plasmid Southern blot hybridisation as a secondary screen. It also allowed the development of a rapid differential library screening protocol with a low false positive rate at the level of filter hybridisation. The majority of the clones picked by the method however, represented cDNAs which although differentially represented between the libraries, were not actually differentially expressed when assayed by *in situ* hybridisation analysis. As has been mentioned, this observation was also made for clones selected as being differential after subtractive hybridisation and differential screening (Dr. S. Dunwoodie and Dr. S. Harrison, personal communication). The method still has potential as a rapid way of differentially screening between libraries; the rate limiting step however is in the isolation of full length clones for further analysis (the differential display products are too small for the direct generation of *in situ* hybridisation probes). In this case, plasmid libraries had to be screened using a sub-pooling and Southern blotting approach; had the libraries been phage derived, this step would have undoubtedly been expedited.

#### **6.2.4 Secondary screening of differential clones**

A robust secondary screen is essential to reliably establish the veracity of the clones derived from differential bands. Classical methods for this have included northern blotting, RNase protection analysis and reverse transcription PCR with primers specific for cloned sequences (Liang and Pardee, 1992; Bauer *et al.*, 1993; Davis *et al.*, 1996; Ariazi and Gould, 1996; Zhao *et al.*, 1996). More esoteric methods have included reverse northern blots and colony lifts (Liang *et al.*, 1995; Zhao *et al.*, 1996). All of these techniques require access to microgram quantities of RNA in order to be successful in a large scale. In this system however, on account of the limited availability of tissue and the information offered by the technique the only really viable approach is *in situ* hybridisation. Only one attempt at using *in situ* hybridisation as a secondary screen in a differential display experiment has been published (Rivas *et al.*, 1997). This used radioactive *in situ* hybridisation to screen the differential clones generated by a comparison of psoriatic and normal skin biopsy RNA. The data were poor and required computer enhancement in order for differential signal to be adequately visualised. The small probe size available to the workers in this case reduced the efficacy of the secondary screen. Such a problem was borne out in the studies performed here when 7.5 *dpc* and 8.5 *dpc* embryo RNAs were compared. Most probes gave no clear signal.

### **6.3 A modified differential display protocol**

The impeding factors in the use of conventional differential display in screening somites and presomitic mesoderm were now clearly defined. Firstly there was a high background of contaminant clones derived from every band (though it seemed that amongst these one was always truly differential); and secondly the clones were too small to generate probes for successful *in situ* hybridisation studies.

#### **6.3.1 The elimination of false clones**

Conventional differential display requires that the primer used in reverse transcription is employed in the PCR in every cycle at low stringency. Consequently, every cDNA that is reverse transcribed will be amplified at least asymmetrically throughout the process and these products, whether visible on autoradiography or not, will litter the gel tracks. Consequently some will be reamplified along with the *bona fide* differential species when they are excised from the gels. This problem is confounded by the reamplification conditions involving two or more cycles at low annealing stringencies

(depending on whether the original display primers or the 5' extended primers were used).

In order to circumvent these problems a modified protocol has been used in which the reverse transcription primer is prevented from taking part in the PCR and so only cDNAs with the arbitrary primers are amplified. This is a simplification of a method proposed previously but which has not been widely adopted (Diachenko *et al.*, 1996a).

The idea is to use longer arbitrary primers in the PCR reactions (24mers in this case), reverse transcription having been carried out with either a conventional dT<sub>24</sub> primer. The use of a short 3' anchored dT primer for reverse transcription as in the conventional protocol would allow for subsetting of the entire cDNA population at the level of reverse transcription and may allow for a larger total number of bands to be generated. However such a primer was not used here as it was thought that the longer arbitrary primers might anneal less frequently to target cDNA and give rise to fewer products *per se* (Welsh *et al.*, 1992). Only a single arbitrary primer was used per reaction in order to simplify subsequent analysis. After two rounds of PCR at low stringency with long annealing times, the stringency of the reaction was increased above a temperature at which the reverse transcription primer could anneal and give rise to reaction products (figure 4.1). It was possible that some inappropriately amplified cDNAs would persist from the first two cycles of PCR but these would not be reamplified as the reamplification reaction was carried out at high stringency throughout with the 24mer primers. A total of 14 bands have been cloned from gels generated by this protocol and each has given rise to a single clone with no evidence of inappropriate clones. A comparison of the conventional protocol of Liang and Pardee and the modified protocol is given in figure 5.1.

Recently, another group at NIMR have started using this modified differential display technique in a *Xenopus* system. They have been comparing the clones they obtain with those derived from a conventional differential display analysis performed with a commercially acquired kit (Gene Hunter Corporation). Although their studies are incomplete, they have found that the new protocol eliminates background clones which they otherwise encounter when using the kit (Mr. S. Kotecha and Dr. T. Mohun, personal communication).

**Figure 5.1**

The salient features of and differences between the conventional differential display protocol of Liang and Pardee and the modified protocol used here to compare dissected mouse somites and presomitic mesoderm.

## Conventional protocol

(Liang and Pardee, 1992)

## Modified protocol

Short anchored  
dT<sub>12</sub>NX primer

First strand  
synthesis

dT<sub>24</sub> primer



Reverse transcription  
primer plus short  
arbitrary 10mer primer.  
40°C annealing (low  
stringency) temperature  
throughout

PCR

Single long arbitrary  
24mer only.  
Initially 2 cycles at  
40°C (low stringency)  
followed by 40 cycles  
at high stringency



Denaturing  
polyacrylamide gel  
electrophoresis

Electrophoresis

Native polyacrylamide  
gel electrophoresis



Low stringency  
throughout using  
differential display  
primers

Reamplification  
of eluted bands

High stringency  
throughout using the  
long arbitrary primer

Figure 5.1

### **6.3.2 The generation of long clones**

The conventional differential display protocol had given rise to bands long enough (> 400bp) for *in situ* hybridisation analysis, but for the reasons detailed earlier, these had proved difficult to reamplify. Using the modified protocol, much larger bands (up to 1.6Kb) could be cleanly reamplified and cloned, in keeping with the predictions of others (Ayala *et al.*, 1995; Diachenko *et al.*, 1996a). This facilitated the direct use of differential display clones generated from a comparison of dissected murine embryonic structures in a secondary screen based on wholemount *in situ* hybridisation.

### **6.3.3 The frequency of true positive bands**

Of seven differential bands (each giving rise to single clones) generated by the modified protocol three gave *in situ* hybridisation patterns consistent with the differential display; the rest gave no signal. No clone specific for the somites gave a signal in the presomitic mesoderm or *vice versa*. More clones would undoubtedly have been generated had a larger number of primers been used in the screen but the veracity of the method with a low false positive rate has been established.

### **6.3.4 Primer annealing sites**

As with the conventional differential display protocol, clones derived from the modified protocol could be generated by the arbitrary primer annealing at any position in the corresponding full length cDNA (where the sequence of the cloned display product was already known). This is what would be intuitively expected of any such randomly primed approach. As before, the 3' most four to six bases of the primer were always exactly homologous to the target. It might be predicted that the longer primers would find targets less frequently in any given cDNA and so the mean product size would be larger than that observed with smaller primers. Tentatively it could be said that this is true from the experience here, though many more gels would have to be run before a more accurate impression can be established.

## **6.4 The abundance of clones isolated by arbitrarily primed PCR**

The question of how abundant the genes that can be detected as differential by differential display and RAP-PCR actually are has received little attention. The majority of biologists who use the technology are primarily interested in the rapid isolation of differential genes. Anecdotal experience from several sources has suggested that rather low abundance species tend to be detected as differential (Conway 1995; Guimarães *et*

*al.*, 1995a, b; Rafaeloff, 1996). As previously outlined, the 'C<sub>0</sub>t effect' (Mathieu-Daudé *et al.*, 1996b), by dint of which the PCR plateaux for amplicons derived from abundant messages are reached before those derived from rarer species, tends to normalise the abundance of products in a differential display. It might be a reasonable assumption therefore that these protocols do detect rare differential species. The most serious attempt to measure how rare a transcript differential display could detect however has suggested that the method could not detect a 4pg spiked contaminant in 2µg total RNA (Bertioli *et al.*, 1995), whereas specific short primers for the spiked RNA could detect a product. This experiment was performed with a limited array of arbitrary primers and it is hard to say that had some other set of primers been used that the contaminant would still not have been detected.

The impression from the experiments shown here is that low abundance species are detected as being differential. The differential species from the library experiments tended to have very low representations in the libraries by plasmid Southern blot hybridisation (figure 3.2; Harrison *et al.*, 1995) and the majority gave rise to probes which provided undetectable signals by *in situ* hybridisation. In the experiments involving somites and presomitic mesoderm, clones that came from differential bands but failed in the secondary screen all gave no *in situ* hybridisation signal. Although this is in no way a formal assessment of the question, it tends to confirm the opinions of others.

## **6.5 A family of transmembrane molecules expressed in the somite**

Strategies such as differential display are designed to isolate differentially expressed genes without any *a priori* information regarding their sequence. The classes of genes isolated depend entirely on the good fortune or otherwise of the experimenter and they may or may not be novel.

The studies here that were comparing presomitic mesoderm and somites revealed the clone S6 to have a very intriguing pattern of expression. It is expressed in a subset of the myotome that is marked by the bHLH transcription factor *Myf-5* and furthermore appears to mark a subset of *Myf-5* expressing cells that is delineated by responsiveness to a particular enhancer of that gene in transgenic mice (Ott *et al.*, 1994; Dr. D. Summerbell, personal communication). In addition, S6 was shown by sequence analysis to contain a leucine rich repeat motif and consequently was thought likely to have close homologues. Because of these two factors, it was decided to study the gene in more detail and attempt to isolate related species which might also have interesting somitic patterns of expression.

### **6.5.1 The leucine rich repeat motif**

The leucine rich repeat (LRR) protein motif was first observed in  $\alpha 2$  - glycoprotein, a component of normal human serum which is up - regulated in certain chronic debilitating diseases (Takahashi *et al.*, 1985; Kobe and Deisenhofer, 1994). Today, it has been identified in many eukaryotic and prokaryotic proteins and which represent a diverse group of molecules with many functions and cellular locations.

The LRR has a consensus sequence consisting predominantly of leucines but also other aliphatic amino acids (reviewed in Kobe and Deisenhofer, 1994). The usual length of an LRR is 24 residues but the repeats may contain as few as 20 or as many as 29. Typically they are present in tandem arrays of upto 30, but can exist as single motifs. All known LRRs have leucine or some other aliphatic amino acid at positions 2, 5, 7, 12, 16, 21 and 24 with asparagine, cysteine, methionine or threonine at position 10. The consensus is quite degenerate and may have been often overlooked by early sequence similarity search algorithms.

LRR motifs are often but not always flanked by cysteine rich motifs at either the amino or carboxyl ends or at both in some cases. These flanking sequences contain four similarly spaced cysteines over a region of about 20 amino acids for the amino - flanking regions and about 50 for the carboxyl flanking regions (Rothberg *et al.*, 1990; Kobe and Deisenhofer, 1994). The sequence motifs in the flanking regions may help to further distinguish sub-families of LRR proteins (Rothberg *et al.*, 1990).

#### **6.5.1.1 Evolutionary relationships of LRR genes**

One of the most extensively studied classes of LRR containing genes, the gonadotrophin receptors contain several similarly sized exons in their LRRs that have homology at the exon - intron junctions. In other LRR genes where this has been studied, the introns often locate at similar positions in the repeats (Kobe and Deisenhofer, 1994). This has led to speculation that the repeat domain arose by exon duplication and shuffling from a single prototype exon that carried the equivalent of a single LRR. It is possible that the duplications occurred separately in each LRR sub-family to give the superfamily seen today or that each sub-family arose after initial duplication events. The variation in the exact sizes of the LRR domains in different families suggest a multifactorial origin for the genes.

### 6.5.1.2 The crystal structures of LRRs

Crystallographic techniques are steadily being applied in the analysis of LRR proteins. The best studied is pancreatic porcine ribonuclease inhibitor (Kobe and Deisenhofer, 1994). This protein has an array of 15 LRRs which do not have cysteine rich flanking sequences. The LRRs alternate between those with 28 and those with 29 residues and each corresponds to a “ $\beta$ - $\alpha$ ” structural unit with a short  $\beta$  - strand and an  $\alpha$  - helix parallel to each other. The overall pattern of the “ $\beta$ - $\alpha$ ” units is to generate a curved nonglobular structure shaped like a horseshoe with the helices on its outer surface. It is unlikely that a single LRR would adopt this conformation and contacts between neighbouring repeats appear to be important in its generation. Further data regarding the structure of other LRR containing proteins is awaited.

### 6.5.1.3 The functions of LRR containing proteins

All the LRR containing proteins known at the present time appear to be involved in protein - protein interactions, and many appear to take part in signal transduction pathways (Kobe and Deisenhofer, 1994). The non - globular secondary structure of the LRR is thought to increase the surface area of LRR containing proteins available for interaction with other species.

The largest group of LRR genes is the adhesive protein family members of which contain a similar 24 residue LRR. They include the small proteoglycans (*e.g.* fibromodulin) which help form the extracellular matrix (Kobe and Deisenhofer, 1994); developmentally important transmembrane genes from *drosophila* (*e.g. slit* and *tartan*) (Rothberg *et al.*, 1988; 1990); and the platelet glycoproteins that form the von - Willebrand Factor (vWF) receptor (reviewed by Roth, 1991). In this latter system, the vWF binding site has been mapped to the LRR domain of platelet glycoprotein 1b $\alpha$  and mutations in this domain lead to the one of the rare hereditary bleeding diathesis, Bernard - Soulier disease.

There are several other groups of LRR genes that are transmembrane receptors *e.g.* the G-protein coupled gonadotrophin receptors. Their LRR domains are of different sizes and within each group there are other conserved sequence motifs (Kobe and Deisenhofer, 1994). CD14 is a macrophage receptor for bound bacterial lipopolysaccharide. Unusually for this class of protein, its structure is almost entirely composed of 8 LRRs and it is GPI-anchored to the cell surface (Kobe and Deisenhofer,

1994). It is thought to interact with other proteins to trigger an intracellular response when it encounters its ligand.

### **6.5.2 Neuronal leucine rich repeat proteins**

S6 has sequence identity with a recently cloned mouse gene, neuronal leucine rich repeat protein 1 (NLRR-1) (Taguchi *et al.*, 1996). This was isolated after an EST clone from a human brain cDNA library that contained leucine rich repeats (EST 06184) was used to screen a commercially available neonatal mouse brain cDNA library (Stratagene). The EST was picked from the database on account of its homology to the *drosophila* gene *slit*, which plays a role in central nervous system development (*vide infra*). A full length 3.7 Kb clone (NLRR-1) was isolated by this screen as well as a 217bp EST named NLRR-2 which encoded an iteration of several related LRRs. Rudimentary analyses of the *in situ* hybridisation expression patterns of these species was performed in sections of 13.5 *dpc* embryos and adult brains. NLRR-1 had wide ranging expression throughout the embryonic nervous system and also in the hippocampus and cerebellum of the adult. No evidence of somitic expression can be seen in the published *in situ* hybridisation images. NLRR-2 also showed expression in the adult brain but no clear embryonic expression could be seen.

The same group subsequently cloned by the same method a third related full length species named NLRR-3 (Taniguchi *et al.*, 1996). Like NLRR-1, this contained 11 LRRs with amino and carboxyl flanking cysteine rich repeats, a 5' signal peptide sequence and a 3' transmembrane domain. Similar rudimentary *in situ* hybridisation studies showed signal in the developing central nervous system (dorsal root ganglia, thalamus and nasal epithelium) of 13.5 *dpc* embryos. No somitic expression could be determined from the published data. In the adult brain expression was restricted to hippocampus, dentate gyrus and olfactory bulb.

Using wholemount *in situ* hybridisation and studying much earlier post - gastrulation embryos, the striking areas of NLRR-1 expression are in the somite and the midline of the forebrain (figure 4.3a). Later embryos than the 30 -35 somite stage (10.5 *dpc*) were not studied as the technique is generally not amenable to these stages. Recently, section data (not presented here) have shown the gene to be expressed in migrating cells of the dorsal myotome.

### **6.5.2.1 The isolation of novel NLRR homologues from the mouse embryo**

A simple protein alignment of the sequences of NLRR-1, NLRR-2 and NLRR-3 obtained from the GenBank database allowed the generation of primers corresponding to conserved sequence motifs (KELG and GDFX) for use in 5' RACE PCR with RNA from somites 5 - 7 of 25 somite stage embryos. Initially this was performed with dATP tailing as no satisfactory upstream consensus could be found. A novel NLRR family member named LRR-4 was isolated and shown to have a classically myotomal pattern of expression (figure 4.9) from the 15 - 20 somite stage. Detailed sequence analysis of LRR-4 suggests that it is the full length clone of the gene represented by the short NLRR-2 EST.

A full alignment of the protein sequence of LRR-4 with those of NLRR-1 and NLRR-3 has allowed for the generation of a primer corresponding to conserved sequences in the amino flanking cysteine rich sequences of these genes. Using this in conjunction with the GDFX primer in PCR of whole embryo (20 - 25 somite stage) RNA which had been reverse transcribed with the KELG primer, two further members of the family have been isolated. One is NLRR-3 and the other has DNA sequence identity with an EST derived from a mouse placenta cDNA library (EST aa020425). In addition to this, the EST database has been searched unsuccessfully using the BLASTn algorithm for further homologues.

### **6.5.2.2 Homologues of the NLRR genes in *drosophila***

The two previously cloned full length NLRR genes and the novel gene LRR-4 presented here all contain iterations of an LRR motif most closely related to those of the *drosophila* genes *slit* and *tartan* (Rothberg *et al.*, 1988; Chang *et al.*, 1993). Both of these genes encode LRR containing molecules important in different cell signaling processes in the embryology of the fly.

*Slit*, which also contains seven EGF repeats at its carboxyl terminus is a secreted protein produced by midline glial cells of the developing embryo. It has four separate tandem arrays of LRRs each of which has its own amino and carboxyl flanking cysteine rich motifs. It appears to play a role in the interactions between midline glial cells, the extracellular matrix and decussating axons (Rothberg *et al.*, 1990).

*Tartan*, on the other hand is a transmembrane molecule with a similar overall structure to the NLRR family. It has a single array of 14 LRRs downstream of its signal

sequence which is flanked by single amino and carboxyl terminal flanking cysteine rich domains. At the carboxyl end, it has a transmembrane domain. Work on this gene in the fly is scant, but it appears to be broadly expressed in the nervous system and in developing muscles. The fly mutants of this gene have disrupted myogenesis amongst their many phenotypic traits (Chang *et al.*, 1993).

Despite their broad structural similarity to *tartan*, the NLRR genes have greater primary sequence homology to the LRRs and flanking domains of *slit*. It must be reiterated however that they are transmembrane molecules and do *not* have EGF repeats.

## **6.6 Conclusions and future directions**

A modified differential display protocol has been developed which can be reliably used to detect transcript differences between single dissected murine embryonic structures using wholemount *in situ* hybridisation as a secondary screen. Using this method, a novel somitic marker which has been recently cloned from the adult mouse brain has been isolated. This is a leucine rich repeat transmembrane protein, NLRR-1 which whilst structurally analogous to *Drosophila* *genetartan*, has greater primary sequence homology over its LRR and flanking domains to *slit*. Using low stringency PCR and 5' RACE, a novel related gene, LRR-4 has been isolated which is expressed in a different domain of the somites.

### **6.6.1 Further applications of differential display**

The protocol presented here can be made to work where only a few cells are available. Consequently large scale dissections are not required and any project can be undertaken with much less effort than is required for related endeavours involving library construction for example. The ability to use *in situ* hybridisation as a secondary screen means that it is possible to compare other dissected organs from the mouse embryo *e.g.* rhombomeres, dissected medial and lateral somites or limb buds. More adventurous applications which might be less guaranteed of a favourable outcome include comparisons of gene expression in cells on either side of a *Hox* gene boundary or comparisons of several individually dissected presomitic mesoderm tissue blocks from different embryos. This might open the possibility of detecting genes whose expression cycles as individual somites are forming.

### **6.6.2 Further studies of the somitic LRR genes**

As mentioned earlier, using low stringency PCR with 5' and 3' degenerate primers, two further members of the NLRR family have been cloned from 25 somite stage mouse embryonic RNA. One of these is the previously known NLRR-3 and the other corresponds to a randomly sequenced EST on the database. *In situ* hybridisation studies of these species are on-going. The search for more members of the family is continuing using a low stringency library screening approach as well as with more aggressive PCR using different primers. More detailed *in situ* hybridisation studies with sectioning are planned in order to delineate precisely which somitic components are marked by each gene. It is hoped to carry out fluorescence *in situ* hybridisation analyses in order to map the exact chromosomal location of each member and ascertain whether they are linked or if they map to the sites of any known murine mutations affecting myogenesis. Ultimately some form of functional study will be required. It is most likely that the two genes presented here play some role in myogenesis. Clearly, targeted mutagenesis of each member of the family both individually and in conjunction with its relatives would be the most complete approach but this is a long term project which will demand considerable resources. It is hoped to gain some information rapidly by either assaying whether any of these genes will rescue the *drosophila tartan* mutation or whether antisense oligonucleotides derived from their appropriate homologues could cause embryological disruption in amphibian or avian systems.

## Acknowledgements

This work was funded by a MRC Clinical Training Fellowship. I am grateful to the Council and to the National Institute for Medical Research for continuing to fund me until a suitable opportunity arose for a to return to clinical practice. I am thankful also to Dr. Peter Rigby for allowing me to work in his group. Despite his fundamental misunderstandings regarding English Premier League football, he remains the smartest person that I have encountered at Mill Hill.

During the course of this work, I have met many people at NIMR who have helped and taught me. I wish to take this opportunity to thank them all for their support. I am especially grateful to my good friend and teacher Dr. Dennis Summerbell, upon whose great technical skills this project was based. His many discourses regarding somitogenesis and Test Cricket, allied to his seemingly insatiable appetite for hot - cross buns will remain a fond memory of my time at the Institute.

At a more personal level, I wish to thank Prof. Lucio Luzzatto, my first clinical boss for introducing me to Peter; Ragini, my sister for her expertise with bibliographic software; and Jaya my wife, who has reconciled herself with considerable fortitude to the fact that I have shunned Harley Street for a career in academic medicine, and without whose love and support I would have forsaken everything.

## Bibliography

- Aasheim H.C., Deggerdal A., Smeland E.B. and Hornes E.** (1994). A simple subtraction method for the isolation of cell-specific genes using magnetic monodisperse polymer particles. *Biotechniques*. **16** 716-21.
- Ayala M., Balint R.F., Fernandez-de-Cossio M.E., Canaan-Haden L., Larrick J.W. and Gavilondo J.V.** (1994) New primer strategy improves the precision of differential display. *Biotechniques*. **18** 841-849.
- Adams M.D., Dubnick M., Kerlavage A.R., Moreno R., Kelley J.M., Utterback T.R., Nagle J.W., Fields C. and Venter J.C.** (1992). Sequence identification of 2,375 human brain genes. *Nature*. **355** 632-4.
- Adams M.D., Soares M.B., Kerlavage A.R., Fields C. and Venter J.C.** (1993). Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. *Nat Genet*. **4** 373-80.
- Ariazi E.A. and Gould M.N.** (1996). Identifying differential gene expression in monoterpene-treated mammary carcinomas using subtractive display. *J Biol Chem*. **271** 29286-94.
- Artavanis Tsakonas S., Matsuno K. and Fortini M.E.** (1995). Notch signaling. *Science*. **268** 225-32.
- Bagnall K.M., Higgins S.J. and Sanders E.J.** (1988). The contribution made by a single somite to the vertebral column: experimental evidence in support of resegmentation using the chick-quail chimaera model. *Development*. **103** 69-85.
- Bauer D., Muller H., Reich J., Riedel H., Ahrenkiel V., Warthoe P. and Strauss M.** (1993). Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR). *Nucleic Acids Res*. **21** 4272-80.
- Bertioli D.J., Schlichter U.H., Adams M.J., Burrows P.R., Steinbiss H.H. and Antoniow J.F.** (1995). An analysis of differential display shows a strong bias towards high copy number mRNAs. *Nucleic Acids Res*. **23** 4520-3.
- Bettenhausen B. and Gossler A.** (1995). Efficient isolation of novel mouse genes differentially expressed in early postimplantation embryos. *Genomics*. **28** 436-41.
- Bettenhausen B., Hrabe de Angelis M., Simon D., Guenet J.L. and Gossler A.** (1995). Transient and restricted expression during mouse embryogenesis of *Dll1*, a murine gene closely related to *Drosophila Delta*. *Development*. **121** 2407-18.
- Bevan S.G., Southey M.C., Armes J.E., Venter D.J. and Newgreen D.F.** (1996). Spatiotemporally exact cDNA libraries from quail embryos: a resource for studying neural crest development and neurocristopathies. *Genomics*. **38** 206-14.
- Bishop J.O., Morton J.G., Rosbash M. and Richardson M.** (1974). Three abundance classes in HeLa cell messenger RNA. *Nature*. **250** 199-204.
- Bonaldo M.F., Lennon G. and Soares M.B.** (1996). Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res*. **6** 791-806.

- Bork P. and Beckmann G.** (1993). The CUB domain. A widespread module in developmentally regulated proteins. *J Mol Biol.* **231** 539-45.
- Bouaboula M., Legoux P., Pessegue B., Delpech B., Dumont X., Piechaczyk M., Casellas P. and Shire D.** (1992). Standardization of mRNA titration using a polymerase chain reaction method involving co-amplification with a multispecific internal control. *J Biol Chem.* **267** 21830-8.
- Brady G. and Iscove N.N.** (1993). Construction of cDNA libraries from single cells. *Methods Enzymol.* **225** 611-23.
- Brenner D.G., Lin-Chao S. and Cohen S.N.** (1989). Analysis of mammalian cell genetic regulation *in situ* by using retrovirus-derived "portable exons" carrying the *Escherichia coli lacZ* gene. *Proc Natl Acad Sci U S A.* **86** 5517-21.
- Burgess R., Cserjesi P., Ligon K.L. and Olson E.N.** (1995). *Paraxis*: a basic helix-loop-helix protein expressed in paraxial mesoderm and developing somites. *Dev Biol.* **168** 296-306.
- Burgess R., Rawls A., Brown D., Bradley A. and Olson E.N.** (1996). Requirement of the *paraxis* gene for somite formation and musculoskeletal patterning. *Nature.* **384** 570-3.
- Chang Z., Price B.D., Bockheim S., Boedigheimer M.J., Smith R. and Laughon A.** (1993). Molecular and genetic characterization of the *Drosophila tartan* gene. *Dev Biol.* **160** 315-32.
- Chee M., Yang R., Hubbell E., Berno A., Huang X.C., Stern D., Winkler J., Lockhart D.J., Morris M.S. and Fodor S.P.** (1996). Accessing genetic information with high-density DNA arrays. *Science.* **274** 610-4.
- Chomczynski P. and Sacchi N.** (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. *Anal Biochem.* **162** 156-9.
- Clegg K.B. and Piko L.** (1983). Poly(A) length, cytoplasmic adenylation and synthesis of poly(A)+ RNA in early mouse embryos. *Dev Biol.* **95** 331-41.
- Conlon F.L., Lyons K.M., Takaesu N., Barth K.S., Kispert A., Herrmann B. and Robertson E.J.** (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development.* **120** 1919-28.
- Conlon R.A., Reaume A.G. and Rossant J.** (1995). *Notch1* is required for the coordinate segmentation of somites. *Development.* **121** 1533-45.
- Conway G.** (1995). A novel gene expressed during zebrafish gastrulation identified by differential RNA display. *Mech Dev.* **52** 383-91.
- Cserjesi P., Brown D., Ligon K.L., Lyons G.E., Copeland N.G., Gilbert D.J., Jenkins N.A. and Olson E.N.** (1995). *Scleraxis*: a basic helix-loop-helix protein that prefigures skeletal formation during mouse embryogenesis. *Development.* **121** 1099-110.
- Davidson E.H. and Britten R.J.** (1979). Regulation of gene expression: possible role of repetitive sequences. *Science.* **204** 1052-9.

- Davis W. Jr, De Sousa P.A. and Schultz R.M.** (1996). Transient expression of translation initiation factor eIF-4C during the 2-cell stage of the preimplantation mouse embryo: identification by mRNA differential display and the role of DNA replication in zygotic gene activation. *Dev Biol.* **174** 190-201.
- Diachenko L.B., Ledesma J., Chenchik A.A. and Siebert P.D.** (1996 a). Combining the technique of RNA fingerprinting and differential display to obtain differentially expressed mRNA. *Biochem-Biophys-Res-Commun.* **219** 824-8.
- Diatchenko L., Lau Y.F., Campbell A.P., Chenchik A., Moqadam F., Huang B., Lukyanov S., Lukyanov K., Gurskaya N., Sverdlov E.D. and Siebert P.D.** (1996 b). Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci U S A.* **93** 6025-30.
- Donohue P.J., Alberts G.F., Hampton B.S. and Winkles J.A.** (1994). A delayed-early gene activated by fibroblast growth factor-1 encodes a protein related to aldose reductase. *J Biol Chem.* **269** 8604-9.
- Donohue P.J., Alberts G.F., Guo Y. and Winkles J.A.** (1995). Identification by targeted differential display of an immediate early gene encoding a putative serine/threonine kinase. *J Biol Chem.* **270** 10351-7.
- Dulac C. and 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** 584-91.
- Eberwine J., Spencer C., Miyashiro K., Mackler S. and Finnell R.** (1992). Complementary DNA synthesis *in situ*: methods and applications. *Methods Enzymol.* **216** 80-100.
- Eberwine J., Yeh H., Miyashiro K., Cao Y., Nair S., Finnell R., Zettel M. and Coleman P.** (1992). Analysis of gene expression in single live neurons. *Proc Natl Acad Sci U S A.* **89** 3010-4.
- Essner J.J., Laing J.G., Beyer E.C., Johnson R.G. and Hackett P.B. Jr** (1996). Expression of zebrafish connexin 43.4 in the notochord and tail bud of wild-type and mutant no tail embryos. *Dev Biol.* **177** 449-62.
- Fischer A., Baum N., Saedler H. and Theissen G.** (1995). Chromosomal mapping of the MADS-box multigene family in *Zea mays* reveals dispersed distribution of allelic genes as well as transposed copies. *Nucleic-Acids-Res.* **23** 1901-11.
- Ghosh S.** (1996). A novel ligation mediated-PCR based strategy for construction of subtraction libraries from limiting amounts of mRNA. *Nucleic Acids Res.* **24** 795-6.
- Greenwald I. and Rubin G.M.** (1992). Making a difference: the role of cell-cell interactions in establishing separate identities for equivalent cells. *Cell.* **68** 271-81.
- Guimarães M.J., Fernando Bazan J., Zlotnik A., Wiles M.V., Grimaldi J.C., Lee F. and McClanahan T.** (1995 a). A new approach to the study of haematopoietic development in the yolk sac and embryoid bodies. *Development* **121** 3335-46

**Guimarães M.J., Lee F., Zlotnik A. and McClanahan T.** (1995 b). Differential display by PCR: novel findings and applications. *Nucleic Acids Res.* **23** 1832-3.

**Gurr S.J., McPherson M.J., Scollan C., Atkinson H.J. and Bowles D.J.** (1991). Gene expression in nematode-infected plant roots. *Mol Gen Genet.* **226** 361-6.

**Hakvoort T.B., Spijkers J.A., Vermeulen J.L. and Lamers W.H.** (1996). Preparation of a differentially expressed, full-length cDNA expression library by RecA-mediated triple-strand formation with subtractively enriched cDNA fragments. *Nucleic Acids Res.* **24** 3478-80.

**Hampson I.N., Pope L., Cowling G.J. and Dexter T.M.** (1992). Chemical cross linking subtraction (CCLS): a new method for the generation of subtractive hybridisation probes. *Nucleic Acids Res.* **20** 2899.

**Hara E., Kato T., Nakada S., Sekiya S. and Oda K.** (1991). Subtractive cDNA cloning using oligo(dT)<sub>30</sub>-latex and PCR: isolation of cDNA clones specific to undifferentiated human embryonal carcinoma cells. *Nucleic Acids Res.* **19** 7097-104.

**Harrison S.M., Dunwoodie S.L., Arkell R.M., Lehrach H. and Beddington R.S.** (1995). Isolation of novel tissue-specific genes from cDNA libraries representing the individual tissue constituents of the gastrulating mouse embryo. *Development.* **121** 2479-89.

**Hill D.P. and Wurst W.** (1993). Screening for novel pattern formation genes using gene trap approaches. *Methods Enzymol.* **225** 664-81.

**Hsu D.K., Donohue P.J., Alberts G.F. and Winkles J.A.** (1993). Fibroblast growth factor-1 induces phosphofructokinase, fatty acid synthase and Ca(2+)-ATPase mRNA expression in NIH 3T3 cells. *Biochem Biophys Res Commun.* **197** 1483-91.

**Hsu D.K.W., Guo Y., Alberts G.F., Copeland N.G., Gilbert D.J., Jenkins N.A., Peifley K.A. and Winkles J.A.** (1996). Identification of a murine TEF-1-related gene expressed after mitogenic stimulation of quiescent fibroblasts and during myogenic differentiation. *J Biol Chem.* **271** 13786-95.

**Hubank M. and Schatz D.G.** (1994). Identifying differences in mRNA expression by representational difference analysis of cDNA. *Nucleic Acids Res.* **22** 5640-8.

**Innis M.A. and Gelfand D.H.** (1990) in "PCR protocols. A guide to methods and applications" Eds. Innis M.A., Gelfand D.H., Sninsky J.J. and White T.J. (Academic Press)

**Ito T., Kito K., Adati N., Mitsui Y., Hagiwara H. and Sakaki Y.** (1994). Fluorescent differential display: arbitrarily primed RT-PCR fingerprinting on an automated DNA sequencer. *FEBS Lett.* **351** 231-6.

**Ivanova N.B. and Belyavsky A.V.** (1995). Identification of differentially expressed genes by restriction endonuclease-based gene expression fingerprinting. *Nucleic-Acids-Res.* **23** 2954-8.

- Jensen J., Serup P., Karlsen C., Nielsen T.F. and Madsen O.D.** (1996). mRNA profiling of rat islet tumors reveals nkx 6.1 as a beta-cell-specific homeodomain transcription factor. *J Biol Chem.* **271** 18749-58.
- Johnson S.W., Lissy N.A., Miller P.D., Testa J.R., Ozols R.F. and Hamilton T.C.** (1996). Identification of zinc finger mRNAs using domain-specific differential display. *Anal Biochem.* **236** 348-52.
- Joshi C.P., Kumar S. and Nguyen H.T.** (1996). Application of modified differential display technique for cloning and sequencing of the 3' region from three putative members of wheat HSP70 gene family. *Plant-Mol-Biol.* **30** 641-6.
- Joyner A.L.** (1991). Gene targeting and gene trap screens using embryonic stem cells: new approaches to mammalian development. *Bioessays.* **13** 649-56.
- Ko M.S.** (1990). An 'equalized cDNA library' by the reassociation of short double-stranded cDNAs. *Nucleic Acids Res.* **18** 5705-11.
- Kobe B. and Deisenhofer J.** (1994). The leucine-rich repeat: a versatile binding motif. *Trends Biochem Sci.* **19** 415-21.
- Koetsier PA., Schorr J., Doerfler W.** (1993). A rapid optimized protocol for downward alkaline Southern blotting of DNA. *Biotechniques.* **15** 260-262
- Kopan R., Nye J.S. and Weintraub H.** (1994). The intracellular domain of mouse Notch: a constitutively activated repressor of myogenesis directed at the basic helix-loop-helix region of MyoD. *Development.* **120** 2385-96.
- Lambert K.N. and Williamson V.M.** (1993). cDNA library construction from small amounts of RNA using paramagnetic beads and PCR. *Nucleic Acids Res.* **21** 775-6.
- Li F., Barnathan E.S. and Kariko K.** (1994). Rapid method for screening and cloning cDNAs generated in differential mRNA display: application of northern blot for affinity capturing of cDNAs. *Nucleic Acids Res.* **22** 1764-5.
- Liang P., Averboukh L., Keyomarsi K., Sager R. and Pardee A.B.** (1992). Differential display and cloning of messenger RNAs from human breast cancer versus mammary epithelial cells. *Cancer-Res.* **52** 6966-8.
- Liang P., Averboukh L. and Pardee A.B.** (1993). Distribution and cloning of eukaryotic mRNAs by means of differential display: refinements and optimization. *Nucleic-Acids-Res.* **21** 3269-75.
- Liang P., Bauer D., Averboukh L., Warthoe P., Rohrwild M., Muller H., Strauss M. and Pardee A.B.** (1995). Analysis of altered gene expression by differential display. *Methods-Enzymol.* **254** 304-21.
- Liang P. and Pardee A.B.** (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction [see comments] *Science.* **257** 967-71.
- Liang P. and Pardee A.B.** (1995). Recent advances in differential display. *Curr-Opin-Immunol.* **7** 274-80.
- Liang P., Zhu W., Zhang X., Guo Z., O'Connell R.P., Averboukh L., Wang F. and Pardee A.B.** (1994). Differential display using one-base anchored oligo-dT primers. *Nucleic-Acids-Res.* **22** 5763-4.

- Lisitsyn N. and Wigler M.** (1995). Representational difference analysis in detection of genetic lesions in cancer. *Methods-Enzymol.* **254** 291-304.
- Lohmann J., Schickle H. and Bosch T.C.** (1995). REN display, a rapid and efficient method for nonradioactive differential display and mRNA isolation. *Biotechniques.* **18** 200-2.
- Luqmani Y.A. and Lymboura M.** (1994). Subtraction hybridization cloning of RNA amplified from different cell populations microdissected from cryostat tissue sections. *Anal Biochem.* **222** 102-9.
- Mackler S.A. and Eberwine J.H.** (1994). Cellular adaptation to opiates alters ion-channel mRNA levels. *Proc Natl Acad Sci U S A.* **91** 385-9.
- Mason C., Lake M., Nebreda A. and Old R.** (1996). A novel MAP kinase phosphatase is localised in the branchial arch region and tail tip of *Xenopus* embryos and is inducible by retinoic acid. *Mech Dev.* **55** 133-44.
- Mathieu-Daude F., Cheng R., Welsh J. and McClelland M.** (1996 a). Screening of differentially amplified cDNA products from RNA arbitrarily primed PCR fingerprints using single strand conformation polymorphism (SSCP) gels. *Nucleic Acids Res.* **24** 1504-7.
- Mathieu-Daude F., Welsh J., Vogt T. and McClelland M.** (1996 b). DNA rehybridization during PCR: the 'Cot effect' and its consequences. *Nucleic Acids Res.* **24** 2080-6.
- McClelland M., Mathieu Daude F. and Welsh J.** (1995). RNA fingerprinting and differential display using arbitrarily primed PCR. *Trends-Genet.* **11** 242-6.
- McClelland M., Ralph D., Cheng R. and Welsh J.** (1994). Interactions among regulators of RNA abundance characterized using RNA fingerprinting by arbitrarily primed PCR. *Nucleic-Acids-Res.* **22** 4419-31.
- Melton D.A., Krieg P.A., Rebagliati M.R., Maniatis T., Zinn K. and Green M.R.** (1984). Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res* **12** 7035-56
- Miyashiro K., Dichter M. and Eberwine J.** (1994). On the nature and differential distribution of mRNAs in hippocampal neurites: implications for neuronal functioning. *Proc Natl Acad Sci U S A.* **91** 10800-4.
- Mohn K.L., Laz T.M., Melby A.E. and Taub R.** (1990). Immediate-early gene expression differs between regenerating liver, insulin-stimulated H-35 cells, and mitogen-stimulated Balb/c 3T3 cells. Liver-specific induction patterns of gene 33, phosphoenolpyruvate carboxykinase, and the jun, fos, and egr families. *J Biol Chem.* **265** 21914-21.
- Murphy N.B. and Pelle R.** (1994). The use of arbitrary primers and the RADES method for the rapid identification of developmentally regulated genes in trypanosomes. *Gene.* **141** 53-61.
- Nedivi E., Hevroni D., Naot D., Israeli D. and Citri Y.** (1993). Numerous candidate plasticity-related genes revealed by differential cDNA cloning. *Nature.* **363** 718-22.

**Nikaido T., Bradley D.W. and Pardee A.B.** (1991). Molecular cloning of transcripts that accumulate during the late G1 phase in cultured mouse cells. *Exp-Cell-Res.* **192** 102-9.

**Nishi M., Kumar N.M. and Gilula N.B.** (1991). Developmental regulation of gap junction gene expression during mouse embryonic development. *Dev Biol.* **146** 117-30.

**Ott M.O., Bober E., Lyons G., Arnold H. and Buckingham M.** (1991). Early expression of the myogenic regulatory gene, *myf-5*, in precursor cells of skeletal muscle in the mouse embryo. *Development.* **111** 1097-107.

**Pourquie O., Fan C.M., Coltey M., Hirsinger E., Watanabe Y., Breant C., Francis West P., Brickell P., Tessier Lavigne M. and Le Douarin N.M.** (1996). Lateral and axial signals involved in avian somite patterning: a role for BMP4. *Cell.* **84** 461-71.

**Prashar Y. and Weissman S.M.** (1996). Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs. *Proc Natl Acad Sci U S A.* **93** 659-63.

**Rafaeloff R., Qin X.F., Barlow S.W., Rosenberg L. and Vinik A.I.** (1996). Identification of differentially expressed genes induced in pancreatic islet neogenesis. *FEBS Lett.* **378** 219-23.

**Ralph D., McClelland M. and Welsh J.** (1993). RNA fingerprinting using arbitrarily primed PCR identifies differentially regulated RNAs in mink lung (Mv1Lu) cells growth arrested by transforming growth factor beta 1. *Proc Natl Acad Sci U S A.* **90** 10710-4.

**Rivas M.V., Jarvis E.D., Morisaki S., Carbonaro H., Gottlieb A.B. and Krueger J.G.** (1997). Identification of aberrantly regulated genes in diseased skin using the cDNA differential display technique. *J Invest Dermatol.* **108** 188-94.

**Roth G.J.** (1991). Developing relationships: arterial platelet adhesion, glycoprotein Ib, and leucine-rich glycoproteins. *Blood.* **77** 5-19.

**Rothberg J.M., Hartley D.A., Walther Z. and Artavanis-Tsakonas S.** (1988). *slit*: an EGF-homologous locus of *D. melanogaster* involved in the development of the embryonic central nervous system. *Cell.* **55** 1047-59.

**Rothberg J.M., Jacobs J.R., Goodman C.S. and Artavanis Tsakonas S.** (1990). *slit*: an extracellular protein necessary for development of midline glia and commissural axon pathways contains both EGF and LRR domains. *Genes Dev.* **4** 2169-87.

**Rothstein J.L., Johnson D., Jessee J., Skowronski J., DeLoia J.A., Solter D. and Knowles B.B.** (1993). Construction of primary and subtracted cDNA libraries from early embryos. *Methods Enzymol.* **225** 587-610.

**Rothstein J.L., Johnson D., DeLoia J.A., Skowronski J., Solter D. and Knowles B.** (1992). Gene expression during preimplantation mouse development. *Genes Dev.* **6** 1190-201.

**Ruangvoravat C.P. and Lo C.W.** (1992). Connexin 43 expression in the mouse embryo: localization of transcripts within developmentally significant domains. *Dev Dyn.* **194** 261-81.

- Rubenstein J.L., Brice A.E., Ciaranello R.D., Denney D., Porteus M.H. and Usdin T.B.** (1990). Subtractive hybridization system using single-stranded phagemids with directional inserts. *Nucleic Acids Res.* **18** 4833-42.
- Sambrook J., Fritsch E.F. and Maniatis T.** (1989) In "Molecular cloning. A laboratory Manual" CSH Laboratory Press
- Sasai Y., Lu B., Steinbeisser H., Geissert D., Gont L.K. and De Robertis E.M.** (1994). *Xenopus chordin*: a novel dorsalizing factor activated by organizer-specific homeobox genes. *Cell.* **79** 779-90.
- Schiller M.R., Mains R.E. and Eipper B.A.** (1995). A neuroendocrine-specific protein localized to the endoplasmic reticulum by distal degradation. *J Biol Chem.* **270** 26129-38.
- Senden N.H., Timmer E.D., Boers J.E., van de Velde H.J., Roebroek A.J., Van de Ven W.J., Broers J.L. and Ramaekers F.C.** (1996). Neuroendocrine-specific protein C (NSP-C): subcellular localization and differential expression in relation to NSP-A. *Eur J Cell Biol.* **69** 197-213.
- Skarnes W.C., Moss J.E., Hurtley S.M. and Beddington R.S.** (1995). Capturing genes encoding membrane and secreted proteins important for mouse development. *Proc Natl Acad Sci U S A.* **92** 6592-6.
- Soares M.B., Bonaldo M.F., Jelene P., Su L., Lawton L. and Efstratiadis A.** (1994). Construction and characterization of a normalized cDNA library. *Proc Natl Acad Sci U S A.* **91** 9228-32.
- Sokolov B.P. and Prockop D.J.** (1994). A rapid and simple PCR-based method for isolation of cDNAs from differentially expressed genes. *Nucleic Acids Res.* **22** 4009-15.
- Sompayrac L., Jane S., Burn T.C., Tenen D.G. and Danna K.J.** (1995). Overcoming limitations of the mRNA differential display technique. *Nucleic Acids Res.* **23** 4738-9.
- Stone B. and Wharton W.** (1994). Targeted RNA fingerprinting: the cloning of differentially-expressed cDNA fragments enriched for members of the zinc finger gene family. *Nucleic Acids Res.* **22** 2612-8.
- Sullivan R., Ruangvoravat C., Joo D., Morgan J., Wang B.L., Wang X.K. and Lo C.W.** (1993). Structure, sequence and expression of the mouse Cx43 gene encoding connexin 43. *Gene.* **130** 191-9.
- Swaroop A., Xu J.Z., Agarwal N. and Weissman S.M.** (1991). A simple and efficient cDNA library subtraction procedure: isolation of human retina-specific cDNA clones. *Nucleic-Acids-Res.* **19** 1954.
- Swiatek P.J., Lindsell C.E., del Amo F.F., Weinmaster G. and Gridley T.** (1994). Notch1 is essential for postimplantation development in mice. *Genes Dev.* **8** 707-19.
- Taguchi A., Wanaka A., Mori T., Matsumoto K., Imai Y., Tagaki T. and Tohyama M.** (1996). Molecular cloning of novel leucine-rich repeat proteins and their expression in the developing mouse nervous system. *Brain Res Mol Brain Res.* **35** 31-40.

- Takahashi N., Takahashi Y. and Putnam F.W.** (1985). Periodicity of leucine and tandem repetition of a 24 - amino acid segment in the primary structure of leucine rich  $\alpha$ 2-glycoprotein of human serum. *Proc Natl Acad Sci U S A.* **82** 1906-10.
- Tam P.P. and Beddington R.S.** (1987). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development.* **99** 109-26.
- Tam P.P. and Trainor P.A.** (1994). Specification and segmentation of the paraxial mesoderm. *Anat Embryol Berl.* **189** 275-305.
- Taniguchi H., Tohyama M. and Takagi T.** (1996). Cloning and expression of a novel gene for a protein with leucine-rich repeats in the developing mouse nervous system. *Brain Res Mol Brain Res.* **36** 45-52.
- Tecott L.H., Barchas J.D. and Eberwine J.H.** (1988). *In situ* transcription: specific synthesis of complementary DNA in fixed tissue sections. *Science.* **240** 1661-4.
- Usui H., Falk J.D., Dopazo A., de Lecea L., Erlander M.G. and Sutcliffe J.G.** (1994). Isolation of clones of rat striatum-specific mRNAs by directional tag PCR subtraction. *J Neurosci.* **14** 4915-26.
- Velculescu V.E., Zhang L., Vogelstein B. and Kinzler K.W.** (1995). Serial analysis of gene expression. *Science.* **270** 484-7.
- Vitek M.P., Kreissman S.G. and Gross R.H.** (1981). The isolation of ecdysterone inducible genes by hybridization subtraction chromatography. *Nucleic Acids Res.* **9** 1191-1202.
- Vogeli Lange R., Burckert N., Boller T. and Wiemken A.** (1996). Rapid selection and classification of positive clones generated by mRNA differential display. *Nucleic Acids Res.* **24** 1385-6.
- Welsh J., Chada K., Dalal S.S., Cheng R., Ralph D. and McClelland M.** (1992). Arbitrarily primed PCR fingerprinting of RNA. *Nucleic-Acids-Res.* **20** 4965-70.
- Welsh J. and McClelland M.** (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **18** 7213-8.
- Welsh J. and McClelland M.** (1992). PCR-amplified length polymorphisms in tRNA intergenic spacers for categorizing staphylococci. *Mol-Microbiol.* **6** 1673-80.
- Welsh J., Petersen C. and McClelland M.** (1991). Polymorphisms generated by arbitrarily primed PCR in the mouse: application to strain identification and genetic mapping. *Nucleic Acids Res.* **19** 303-6.
- Wilkinson D.G. and Green J.** (1990) in "Post implantation mammalian embryos. A practical approach." Eds. Copp A.J. and Cockroft D.L. IRL Press
- Williams R., Lendahl U. and Lardelli M.** (1995). Complementary and combinatorial patterns of Notch gene family expression during early mouse development. *Mech Dev.* **53** 357-68.

**Yamaguchi T.P., Harpal K., Henkemeyer M. and Rossant J. (1994).** *fgfr* - 1 is required for embryonic growth and mesodermal patterning during mouse gastrulation. *Genes Dev.* **8** 3032-44.

**Zeng J., Gorski R.A. and Hamer D. (1994).** Differential cDNA cloning by enzymatic degrading subtraction (EDS). *Nucleic-Acids-Res.* **22** 4381-5.

**Zhao S., Ooi S.L., Yang F.C. and Pardee A.B. (1996).** Three methods for identification of true positive cloned cDNA fragment in differential display. *Biotechniques.* **20** 400-4.

**Zimmermann J.W. and Schultz R.M. (1994).** Analysis of gene expression in the preimplantation mouse embryo: use of mRNA differential display *Proc Natl Acad Sci U S A.* **91** 5456-60.