Development of a Resource for the Exploration of Gene Expression in the Mouse Foetus

Elizabeth Ann Campbell CBiol. MIBiol.

A dissertation submitted to the Open University
for the degree of Master of Philosophy

January 2004

Sponsoring and collaborating establishments:

The Sanger Institute,
Wellcome Trust Genome Campus,
Hinxton,
Cambridgeshire,
CB10 1SA

Anatomy Department,
Cambridge University,
Downing Site,
Cambridge,
CB2 3DY

Funded by Parke Davis and The Wellcome Trust

CONTENTS

		Page Number
Acknowledgements		1
ABSTRACT		2
INTRODUCTION		3
•	From gene sequence to gene expression	3
•	A model organism	4
•	Ways to map gene expression patterns – strengths and weaknesses	4
•	Summary of Strengths and Weaknesses of these Gene Expression Methods	12
•	My approach	12
•	The Family of Sox Genes	14
•	Sry Gene	14
•	Figure 1: The Jigsaw Puzzle of Vertebrate Sex Determination	17
•	Sox Family of Genes	18
Figure 2	Protein-DNA interactions in the SRY_HMG/DNA complex	19
Figure 3	Schematic representation of SOX proteins	21
Figure 4	Proposed Phylogeny of the SOX Transcription Factors	22
Figure 5	Sequence alignment of the HMG domains of the family of Sox genes	23
•	Objectives of My Research	24
References		25

CHAPTER 1 Construction of the Panel

Introduction		29
1.1.	Special Equipment and Suppliers	29
1.2.	Materials and Solutions	30
1.3.	Methods	35
1.3.1.	Preparation and maintenance of Animals	35
1.3.2.	Tissue Collection	35
Figure 6	Mouse handling	36
Figure 7	Dissection of mice	36
1.3.3.	Dissection Procedure	37

1.3.4.	Extraction of RNA from Tissues using Bioline Rnace Kit	38
1.3.5.	Deoxyribonuclease treatment	39
1.3.6.	Reverse Transcription	40
1.3.7.	Stock and working solutions	41
1.3.8.	Polymerase Chain Reaction	41
1.3.9.	Typical Cycling Programme	41
1.3.10	Primer design	42
Figure 8	Primer sequences	43
1.4.	Construction of the Panel	44
1.4.1.	Tissue dissection	44
Figure 9	Table of Tissues Dissected	44
1.4.2.	RNA extraction and assessment	45
Figure 10	RNA extraction from 15.5d foetus	45
Figure 11	Agarose gel of RNA	46
Figure 12	Illustration of differing qualities of RNA	47
1.4.3.	Reverse transcription and PCR amplification	48
1.4.4.	Storage of Panel	49
Figure 13	Final format for The Mouse Foetal cDNA Panel	50
1.4.5.	Presecreen	51
Figure 14	Prescreen	51
1.5.	Panel Validation	52
1.6.	Concluding Remarks	54
Figure 15	Genomic contamination – mouse foetal panel	56
Figure 16	Housekeeping genes – mouse foetal panel	57
Figure 17	Sucrase isomaltase – mouse foetal panel	58
Figure 18	Sucrase isomaltase – mouse adult panel	59
Figure 19	Casein, alpha – mouse foetal panel	60
Figure 20	Casein, alpha – mouse adult panel	61
Figure 21	Solute Carrier family 17, Na/H exchanger – mouse foetal panel	62
Figure 22	Solute Carrier family 17, Na/H exchanger – mouse adult panel	63
Figure 23	Calbindin 28K – mouse foetal panel	64

Figure 24 Calbindin 28K – mouse adult panel 65
Figure 25 Fatty Acid Binding Protein – mouse foetal panel 66
Figure 26 Fatty Acid Binding Protein – mouse adult panel 67
Figure 27 Adora 1 – mouse foetal panel 68
Figure 28 Adora 1 – mouse adult panel 69
Figure 29 Hox 5b – mouse foetal panel 70
Figure 30 Hox 5b – mouse adult panel 71
Figure 31 Graph of Gel images of Figure 16 73
Figure 35 Tissue identification of graph position 74
References 75

CHAPTER 2 Profiling the Family of Sox Genes by PCR
Introduction 76
2.1 Special Equipment and Suppliers 76
2.2 Material and Solutions 76
2.3. Establishment of the Sox-related sequences for study 76
Figure 32 Blixem Software 78
Figure 33 Expected amplicon sizes for PCR 79
Figure 34 Primer Sequences 80
2.4.1 Results of the PCR for the Sox Genes 81
Figure 35 Tissue ID 82
Figure 36 Expression Profile of *Sox1* in Mouse Foetal Panel 83
Figure 37 Graphical representation of gel image in Figure 36. 84
Figure 38 Expression Profile of *Sox2* in Mouse Foetal Panel 86
Figure 39 Graphical representation of gel image in Figure 38 87
Figure 40 Expression Profile of *Sox4* in Mouse Foetal Panel 88
Figure 41 Graphical representation of gel image in Figure 40 89
Figure 42 Expression Profile of *Sox5* in Mouse Foetal Panel 90
Figure 43 Expression Profile of *Sox6* in Mouse Foetal Panel 91
Figure 44 Expression Profile of *Sox7* in Mouse Foetal Panel 94

Figure 45 Expression Profile of *Sox8* in Mouse Foetal Panel 95
Figure 46 Expression Profile of *Sox9* in Mouse Foetal Panel 96
Figure 47 Graphical representation of gel image in Figure 46 97
Figure 48 Expression Profile of *Sox10* in Mouse Foetal Panel 98
Figure 49 Expression Profile of *Sox11* in Mouse Foetal Panel 100
Figure 50 Expression Profile of *Sox12* in Mouse Foetal Panel 101
Figure 51 Expression Profile of *Sox13* in Mouse Foetal Panel 104
Figure 52 Expression Profile of *Sox14* in Mouse Foetal Panel 105
Figure 53 Expression Profile of *Sox15* in Mouse Foetal Panel 106
Figure 54 Graphical representation of gel image in Figure 53 107
Figure 55 Expression Profile of *Sox16* in Mouse Foetal Panel 109
Figure 56 Expression Profile of *Sox17* in Mouse Foetal Panel 110
Figure 57 Expression Profile of *Sox19* in Mouse Foetal Panel 112
Figure 58 Expression Profile of *SoxLZ* in Mouse Foetal Panel 113
Figure 59 Expression Profile of *Slc17a2* in Mouse Foetal Panel 114
Figure 60 Expression Profile of *Csna1* in Mouse Foetal Panel 116
Figure 61 Expression Profile of *FABPi* in Mouse Foetal Panel 117
2.4.2 Concluding Remarks 118
References 119

CHAPTER 3 Validation of the Sox Genes PCR data by *in-situ* hybridisation
Introduction 123
3.1 Special Equipment and Suppliers 123
3.2. Materials and Solutions 125
3.3. Methods 135
3.3.1. Preparation of Glass Slides 135
3.3.2. Collection of Tissues 136
3.3.2.1. Sectioning of Tissues 137
3.3.2.2. Fixation of sections 138
3.3.2.3. Storage of fixed sections 138

3.3.3.	Design, synthesis and purification of oligonucleotide probes	139
Figure 62	*In-situ* Probe Sequences	141
3.3.4.	Radioactive labelling and purification of oligonucleotide probes.	142
3.3.4.1.	Radioactive ^{35}S-labelling of oligonucleotide probes.	142
3.3.4.2.	Purification of Probe	143
3.3.4.3.	Removal of unincorporated label	144
Figure 63	Table of label incorporated into Probes	145
3.3.4.4.	Incubation of sections with labelled probe	145
3.3.5.	Post-hybridization treatments	146
3.3.5.1.	Slide washing	146
3.3.6.	Autoradiography	147
3.3.6.1.	Exposure to X-ray film	147
3.3.6.2.	Development of the X-ray film	148
3.3.7.	Emulsion autoradiography	148
3.3.7.1.	Development of emulsion-coated sections	150
3.3.8	Counterstaining	151
3.3.9.	Protocol for capturing images	152
3.3.10.	Brief summary of *In-situ* Protocol	155
Figure 64	Composite of Autoradiograph images of slides	156
Figure 65	Illustration of embryo orientation	157
3.4.	Results	158
3.4.1.	Visual Analysis of Slides	158
Figure 66	Artifacts of silver staining	161
Figure 67	Silver staining Artifacts	162
Figure 68	Table of *Sox2 In-situ* hybridisation	163
Figure 69	Table of *Sox6 In-situ* hybridisation	164
Figure 70	13.5dpc region of spinal cord, hybridisation results (x10)	168
Figure 71	13.5dpc ependymal layer of spinal cord, hybridised to *Sox2* (x40)	169
Figure 72	Illustration of a transverse section of a mouse embryo at 14 days gestation	170
Figure 73	Lung and Snout regions - specific *Sox2* hybridisation to bronchi and primordial vibrissae follicle	171

Figure 74 15.5dpc olfactory region – specific *Sox2, Sox6, Sox16* and *SoxLZ* hybridisation 172
Figure 75 15.5dpc Cochlea hybridised to *Sox2* (x40*)* *173*
Figure 76 17.5dpc, Eye, hybridisation results (x10) 174
Figure 77 17.5 dpc Eye identification 175
Figure 78 Table of specific silver signalling in the eye 177
Figure 79 Neural tissues of 15.5dpc hybridised to *Sox2* 176
3.4.2. Digital Analysis of Slides 178
Figure 80 Data and Graph of 13.5dpc spinal cord silver grain counts 178
3.4.3. Concluding Remarks 181
References 184

CHAPTER 4 Discussion
Introduction 186
4.1. Main Aims 186
4.2. Exploring the Methodology and underlying Theoretical approach 186
4.3. Panel Validation 190
4.4. Profiling the Sox family 191
4.5. Validation of the RTPCR 191
4.6. Refinements of the Strategy 192
4.7. Future uses of this resource 195
References 197

Appendix 1 Images of dissected mouse foetuses 199
Appendix 2 Foetal development 200
Appendix 3 Delta-like and Gtl2 are reciprocally expressed, differentially methylated linked imprinted genes on mouse chromosome 12 Takada, S., Tevendale, M., Baker, J., Georgiades, P., **Campbell, E.,** Freeman, T., Johnson, M. H., Paulsen, M., Ferguson-Smith, A. C. Current Biology, 2000, 10, (18), p1135-1138

Abbreviations

European Molecular Biology Laboratory	EMBL
Expressed sequence tag	EST
sequence tagged site	sts
bacterial artificial chromosome	bac
days post conception	dpc
final concentration	*fc*
catalogue number	#
molar	M
litre	L
millilitre	ml
microlitre	□l
milligram	mg
microgram	□g
nanogram	ng
hour	hr
minutes	mins
seconds	secs
revolutions per minute	rpm
hydrochloric acid	HCl
diethylpyrocarbonate	depc
ethylene diamine tetraacetic acid	EDTA

trizma base	Tris
1 M Tris 0.1 M EDTA	T0.1E
tris borate EDTA buffer	TBE
ethidium bromide	EtBr
deoxyribonuclease 1	Dnase1
moloney murine leukemia virus reverse transcriptase	M-MLV
terminal deoxynucleotidyl transferase	TdT
2’-deoxyribonucleoside 5’-triphosphates	dNTPs
dithiothreitol	DTT
reverse transcription polymerase chain reaction	RTPCR
reverse transcription	RT
complementary deoxynucleic acid	cDNA
ribonucleic acid	RNA
sodium dodecyl sulphate	SDS
sodium chloride	NaCl
bovine serum albumin	BSA
disodium hydrogen orthophosphate	Na_2HPO_4
sodium dihydrogen orthophosphate	NaH_2PO_4
poly acrylamide gel electrophoresis	PAGE

Acknowledgements

I am indebted to Tom Freeman for giving me the impetus to embark on this project. From the Sanger Institute, I would like to thank the many scientists for sharing the fun of science, in particular Kate Rice, Adam Butler and Sarah Hunt for helping to me to understand bioinformatics and Dave Vetrie for his meticulous proof reading.

Of the scientists at Parke Davis, I would like to thank Rob Pinnock, Peter Cox, Kevin Lee and Alistair Dixon for all their support whilst working in their laboratory.

And for the continued and persistent encouragement throughout the course of this work, I owe Jane Rodgers a huge thank you.

I would like to thank Martin Johnson, Anne Ferguson-Smith, Christy Starr and Martin George for supplying the materials and expertise to pursue this study and Peter Wooding for assisting me with the technique of *in-situ* hybridisation. I would also like to thank John Bashford, Adrian Newman and Ian Bolton of the AVMG, for assistance with computing, microscopy and printing of this thesis. And I especially would like to thank Josie McConnell for the varied and valuable discussions.

I am particularly grateful to Martin Johnson for his gentle but untiring encouragement and guidance during this project and for helping me to order my thoughts sufficiently for writing.

And finally I would like to thank two of the most important men on my life, my husband, John and my son, Simon for their patience, humour and support throughout the duration of this work.

Abstract

Seventy five different foetal and placental tissues/organs, from naturally mated C57BL/6J mice, were dissected at 8.5d, 9.5d, 10.5d, 11.5d, 12.5d, 13.5d, 15.5d and17.5 days post-coitum. Tissues were stored at –80^{o}C and ribonucleic acid (RNA) extracted. Good quality RNA, as assessed by OD and agarose gel electrophoresis, was treated with Dnase1 and 50 g was reverse transcribed. The resulting complementary desoxyribonucliec acid (cDNA) was checked for the presence of genomic DNA, and amounts of cDNA adjusted so that equivalent amounts of PCR product from a set of two presumptive housekeeping genes were found in each of the 89 samples. The final panel consisted of cDNA from 75 mouse foetal tissues, 11 adult mouse tissues, 2 glycogen samples and rat, human and mouse genomic DNA, stored in a deep well microtitre plate.

Primers designed to 7 genes with a known expression pattern in adult tissues were used in the polymerase chain reaction to validate the integrity of the panel contents. The usefulness of the panel was tested by running PCR reactions with primers designed to the newly emerging family of *Sox* genes. Expression profiles for 17 members of the *Sox* family were obtained. PCR expression patterns of two of the *Sox* genes (*Sox2* and *6*) were compared with profiles obtained by *in-situ* hybridisation on staged sections of the mouse foetus. A reflection on the process of the panel formation and testing has highlighted a number of refinements to the process of exploring gene expression in fetal and placental tissues.

INTRODUCTION

From gene sequence to gene expression

The physical map of the mouse genomic sequence has been published [3], the basic descriptive human genome project is now completed [4], and the analysis of variation in both genomes is underway [5]. The next big challenge is to locate regions of gene expression and to understand how their expression is controlled and what their expression means mechanistically for development. During foetal developmental, certain genes and pathways are activated in a defined pattern and sequence for a fully functioning organism to develop. Many genes are known to be important during embryogenesis, because, when they are inactivated or only present on one chromosome, developmental pathology can result [6]. Mapping genes that are expressed in a particular tissue at a given time point will help in our understanding of gene hierarchy and function. Moreover, the precise form of gene expression observed is important, as evidenced by gene isoforms *Sox17* [7], *Plectin* [8] and *WT1* [9] performing different functions. Thus, observing either the expression pattern or the sequence information in isolation can be misleading. Genes switched on during development are often developmentally specific, as seen with the Cystic Fibrosis Transmembrane conductance Regulator (*CFTR*) gene [10], but the same gene can also be expressed later in development, often in a variant form through splicing, initiation or post-translational differences. Having an expression map will provide a useful database for intelligently manipulating development, so that cause and effect can be explored.

A model organism

The very early stages of development are comparable in all eutherian mammalian species and events that subsequently occur have many similarities between the species [11]. Beginning with the fusion of the male and female gametes, through pre-implantation and post-implantation stages of embryogenesis to the formation of the embryo and then the foetus, the construction of organs and tissues seems to be broadly conserved at molecular and organisational levels. The mouse is an ideal organism to study the expression of genes during foetal development, as it has been the organism of choice for many genetic, teratogenic, manipulative and developmental studies. Thus, information on expression patterns in the mouse provides a useful guide to patterns in other mammalian species.

Ways to map gene expression patterns – strengths and weaknesses

There are many different ways to study gene expression, including the techniques of northern blotting, nuclease S1 mapping, the ribonuclease protection assay, RTPCR (Reverse Transcription Polymerase Chain Reaction), *in-situ* hybridisation, computational studies, and more recently microarray technologies. Each of these will be described in turn, together with their strengths and weaknesses.

The northern blot has traditionally been applied to measure the total size/mobility of a transcript. It is thus useful for detecting variants due to differential splicing and initiation sites. It can also be used to screen for family gene members from related organisms, by varying the temperature of hybridisation. The northern blot detects

relatively high abundance messages like actin, which make up as much as 0.1% (300pg from 5mg tissue) of the mass of total RNA whereas specific rarer messages exist at levels below 0.001% (10fg – 10pg from 5mg tissue) of the total mass and can be more difficult to distinguish. With care, the northern blot can detect levels as low as 10 pg (picogram) amounts. RNA samples to be studied are first separated on an agarose gel under denaturing conditions, and then transferred to a filter and immobilized. The filter is then hybridized with a labelled probe (usually cDNA - complementary Deoxyribo Nucleci Acid - or RNA) and specific targets are identified by autoradiography or nonradioactive methods such as digoxigenin. The main advantage is that the filter can be screened repeatedly under a variety of stringencies with different probes. Sequences with only partial homology (cDNA from different species or genomic DNA fragments that contain an intron) can be used as probes to identify transcript size and recognize alternatively spliced transcripts. However, the technique requires very high quality, full length RNA at the high concentration of 1mg/ml. Poor quality RNA will result in loss of signal, whilst contributing to the background 'noise', for example, a single nick in 20% of a 4 Kb transcript will reduce the resulting signal by a full 20% (http://www.ambion.com/techlib/basics/northerns/index.html).

Nuclease S1 mapping can be used to identify particular sequences of RNA and can also be used to position cap sites or splice junctions. In this procedure, single stranded DNA probes are generated that are labelled at either the 5' or 3' end. Probes are hybridized to the RNA and single strands digested with single-strand specific nucleases. The residual double strands are separated by PAGE (polyacrylamide gel electrophoresis)

and visualised by autoradiography. The major drawback of the nuclease assay is the difficulty in controlling non-specific digestion of AT rich regions, which often transiently become single stranded at the operating temperature of the nucleases at 16^{o}C and so become available for digestion resulting in inaccurate data [12].

The ribonuclease protection assay, which follows on from nuclease mapping, uses a radiolabelled antisense probe generated to the RNA under investigation. This probe is hybridized to the RNA sample, and the resulting hybrid is treated with single-stranded ribonucleases. The surviving duplex is visualised by PAGE and autoradiography. High specific activity probes are used for rarer messages and low specific activity probes for abundant messages. This method is very sensitive, shows fewer degradation problems and a lower background than the S1 mapping assay. However the northern blot and RTPCR are both better at detecting multi-gene families and inter-species variants.

The technique of RTPCR has been extensively studied and found able to provide a reliable and reproducible method for the global study of gene expression. This technique relies on the efficient and faithful copying of an mRNA strand into a complementary DNA (cDNA) strand and the amplification of specific regions as directed by primers designed to particular regions of the sequence. When conditions are optimal, a single copy of cDNA can be detected in a complex mixture with this approach and it is the method of choice when looking for rare transcripts. Through carefully optimising the reaction conditions, a level of relativity can be achieved whereby within a series of reactions it is possible to illustrate which tissue source has more mRNA than others in the

same series for a given sequence. The main disadvantage is that PCR will identify sequences in genomic DNA as readily as cDNA. Therefore, often it is necessary to DNase 1 treat the RNA prior to reverse transcription, which may damage RNA species. The only effective way to distinguish between the genomic and cDNA contributions is the careful design of primers such that the resulting amplicons differ in size. This can only be done for sequences that contain at least one intron. Quantitative RTPCR is a development of the PCR technique, which makes use of a fluorescent reporter molecule that is cleaved during the extension stage of the PCR cycle. An extension of this technique uses the fluorescent marker sybr green: this marker binds to the double stranded product of PCR, fluorescing significantly more when bound to this double stranded product than in the unbound state. The technique of real time PCR is a more sensitive technique in the search for rare transcripts. Some detection systems can distinguish between different reporter molecules at these low quantities, making it possible to multiplex for different genes in a single sample. However, this is a relatively new technology, and to conduct real-time RTPCR requires expensive instrumentation and reagents.

In-situ hybridisation is a technique that detects specific nucleic acids in morphologically preserved sections. It is a technique that has evolved from immuno-histochemistry, which highlights the morphology of a section with specific stains and antibodies to proteins. Traditionally, tissue sections are used, but whole cells or embryos and chromosome spreads can also be used. The method requires the fixing of samples either to a slide or as a whole mount in a chamber. The specifically designed nucleic acid

is labelled and hybridised to the fixed sample. A wide variety of non-radioactive labels are available, but radioactivity is still regarded as the most sensitive method for the detection of rare sequences. The sequence for labelling is either ordered from oligo synthesising companies and then enzymatically labelled, or PCRed, cloned and transcribed as a radioactive riboprobe. In tissue sections, this technique microscopically locates the region of interest to the internal composition of individual cells. Pathbase (www.pathbase.net) is a database repository of histopathology photomicrographs and macroscopic images of the mouse throughout development, covering all strains, mutant, chimaeric, transgenic and knockout specimens. A similar database to map the protein expression in mouse sections – Atlas project - is in the planning stages at the Sanger Institute. In Edinburgh, a database is being built website (http://genex.hgu.mrc.ac.uk/ (EMAGE)) describing ISH expression patterns in the mouse embryo, which is publicly available. A controlled vocabulary is used during data/image entry to facilitate the linking with related databases. Databases of this nature are likely to become invaluable sources of information to the mouse scientific community.

Computational analysis of the genome has provided the biologist with a wealth of information relating to structure and composition for a number of organisms. In particular it has highlighted the similarities between species and the usefulness of the mouse as a model organism [13]. Of the 30,000 or so genes in the mouse sequence, 99% have direct counterparts in the human sequence [14]. There is 40% direct alignment between mouse and human, with 80% of human genes having one corresponding gene in the mouse. The mouse genome is, however, 14% smaller at 2.5Gb, compared to 2.9Gb

for the human genome. Computer analyses have identified blocks of synteny, between the genomes, mapping the relevant regions with corresponding linking threads to illustrate those areas of synteny. An example of conserved synteny can be found in human chromosome 20: consisting of only three segments, which are identical to regions of mouse chromosome 2, with only one small segment altered in its order on the chromosome [5]. Through linking databases, information regarding sections of the mouse sequence can now be quickly compiled on screen for a thorough evaluation of genes of interest. With the complete annotation of the sequence, that reveals sequences flanking genes to identify promoter elements, non-protein coding transcripts, intron-exon structure, splice variants, alternative polyadenylation sites, sense and antisense pairings, and related genes, the task of the scientific researcher will be made much simpler. From a recent evaluation of the mouse gene Sox 8, using Compugen's Gencarta software, 9 possible transcripts were identified. On closer examination, only 4 of these were thought to be transcribed. From studies of this nature, experiments will in future be more precisely designed to study the biological condition under investigation.

Microarray is a widely available technology that is being employed to address increasingly complex scientific questions. The basic concept behind all microarray experimentation is the precise positioning of probes at high density on a solid support. The probes then act as molecular detectors [15]. From minuscule amounts of starting RNA, from as low as 30ng [16], the expression patterns of thousands of genes can be profiled in a single experiment. The aim is to describe quantitatively the gene expression profile characteristics associated with, for example, particular physiological processes or

behaviours. By comparing patterns under different conditions, the association of particular gene expression profiles with particular types of activity can be quantified. Microarray probes are traditionally nucleic acid sequences as either synthesised oligonucleotides or PCR fragments. Protein arrays, which are composed of, either antibodies, aptamers, small-molecule drugs or phage particles, are currently under development. The target sample is labelled with radioactivity or a range of fluorescent molecules and hybridized to the probe array under stringent conditions. The resulting image is scanned and the location of successful hybridizations recorded together with the intensity of signal. In essence, by using different sources of target molecules, one can map changing temporal profiles or responses of tissues to signals or disease pathogenesis. Thus, this technique has scope in the fields of gene screening, target identification, pathway mapping, developmental biology and disease progression and diagnostic characteristics. This is a powerful technique as illustrated by Miki et al. [17], who profiled cDNA from 49 different mouse tissues and revealed related patterns of expression in related tissues.

For the pharmaceutical industries the microarray approach is beginning to be applied to RNA and protein alterations in early drug screening and non-clinical toxicology studies. For medicine it may become useful in the field of diagnostic biomarkers and patient tailored therapies (pharmacogenetics) [18]. However, initially, this technology requires more stringent validation to be useful to the medical profession. The entire process requires accurate selection, amplification and location of probes, accurate reference sequence information, identification of unique probe oligonucleotides, accurate distinction among multiple products of a single gene, accurate reconstruction of

expressed sample nucleotide sequences, precise image scanning, and reproducible and accurate transformation of image files to numerical data. To be reliable, the probes must hybridize with high sensitivity and specificity, reproducibly between experiments and between laboratories, with a biologically meaningful outcome. One concern is that statistically a single microarray of 10,000 elements, with 99% accuracy may generate as many as 100 false positives, which can seriously affect results. Another source of error is that 50% of expressed eukaryotic genes are estimated to be expressed as splice variants, therefore the precise biological outcome may not be uniquely identifiable.

The enormous amount of data points generated by a single microarray experiment necessitates the application of bioinformatics to analyse results. The Gene Ontology consortium is creating standard contextual terms that are recognisable by computers to aid the analysis of these large data sets. With multiple experiments, data storage becomes an issue. There are a number of database repositories for microarray data, where data sets can be up- and down-loaded for comparative analysis [19]. This technique also requires some expensive instrumentation and significant computational skills and capabilities. A limited number of results from analysed microarray data can be validated with other technologies, such as real time RTPCR or in situ hybridisation. However, it would be unrealistic to attempt the validation in this way of the thousands of data points this technology generates.

Summary of Strengths and Weaknesses of these Gene Expression Methods

Method	Total RNA amounts	Mode of detection	Weakness	Strength
Northern Blot	10 pg – 30 □g	Radiography or nonradiography	Requires very high quality RNA.	Measures relative RNA levels within a single blot.
Nuclease S1 mapping	10pg – 100 □g	Radiography or nonradiography	Non specific digestion of AT rich regions	More sensitive than Northerns, used to map mRNA termini and intron/exon junctions.
Ribonuclease protection	5 femtograms – 100 □g	Radiography or nonradiography	Must know complete sequence of mRNA for probe design.	Up to 10 probes per sample for comparisons of muti-gene families.
Reverse Transcription Polymerase Chain Reaction	1 attog – 40 ng	Ethidium Bromide or Fluorescence	Primer design	Most sensitive method for mRNA detection and quantitation. Tolerates slightly degraded mRNA.
In-Situ hybridization	10-20 copies	Radiography or nonradiography	Long procedure	Precisely localises position of expression.
Computational Studies	N/A	World wide web	Relies on other people's data	Requires validation with biological samples.
Microarray technologies	5 – 100 □g	Radiography or fluorescence	Methods vary between labs. Optimisations ongoing.	Thousands of targets per sample.

My approach

To capitalise on the technological advances in high throughput screening (HTS) for gene expression profiling in as broad a number of tissues as possible, the most effective and sensitive method is the RTPCR with confirmative evidence from *in-situ* hybridization. This approach minimises costs and radioactivity usage. Computational

analysis is essential for collecting and analysing sequence information, for primer/probe design and collating, presenting and evaluating laboratory results.

An earlier study, systematically characterising genes in the adult mouse over a range of tissues, used the technique of RTPCR. This database was submitted electronically to the Jackson Laboratories [20]. Our group had amassed over a thousand PCR primer pairs and had begun to create similar cDNA resources for mouse mammary gland differentiation during pregnancy, transgenic mouse gut [21], cancerous disease states in the human, and drug treatment in rats (much of this work is unpublished due to commercially orientated funding). Finely dissected tissues were supplied by collaborators and processed into cDNA resources to probe for genes of interest in a high through put fashion. With this background, and using the same rationale, I decided to create and test a similar resource of foetal tissues during progression to full term to help identify developmentally significant genes. Prior to joining the Sanger Centre, I had been working on the pre-implantation mouse embryo at the Anatomy Department, Cambridge University and understood the difficulties associated with obtaining reliable data from such small amounts of material. Through advances in HTS and molecular biology, it is becoming increasingly possible to obtain reproducible gene information from these small quantities of starting material.

Profiles of gene expression from this resource will illustrate when and to what degree (relative to other stages/tissues), genes are being expressed during foetal development, using a single method across staged tissues. This knowledge can be useful when studying individual tissues/organs during development, as it is possible to establish correlations, which illustrate which other tissues and organs are being affected by the

same or similar genes, during the time of interest. With the careful design of primers, the approach will demonstrate the presence of mRNA in differing forms (splice variants). In a single experiment, genes expressed exclusively in the foetus (or adult) can be illustrated by using this collection of cDNAs (as shown by [22]). The resource can then be made available to the mouse community with more specific interests.

Having established the resource, we chose to investigate the usefulness of this resource by profiling the family of Sox genes.

The Family of Sox Genes

The Sox (Sry box) gene family, derive their name from the founding gene member Sry (Sex-determining Region of the Y chromosome). These genes are thought to encode transcription factors. However, there are many members of this expanding gene family and their functional role has yet to be fully elucidated. Sox genes are involved in a number of diverse functions, governing cell fate and organ development at different times and in different tissue locations during embryology. They thus provide a good target for trialling this cDNA resource.

Sry Gene

The regulation of sex is controlled genetically in eutherians. The presence of a Y chromosome directs construction of a male gonad from the gonadal rudiment (Ford et al 1958 Nature **181**, p.1565). The later hormonal output of the male gonad then secondarily directs the developing embryo to form a male. The search for the testis-determining

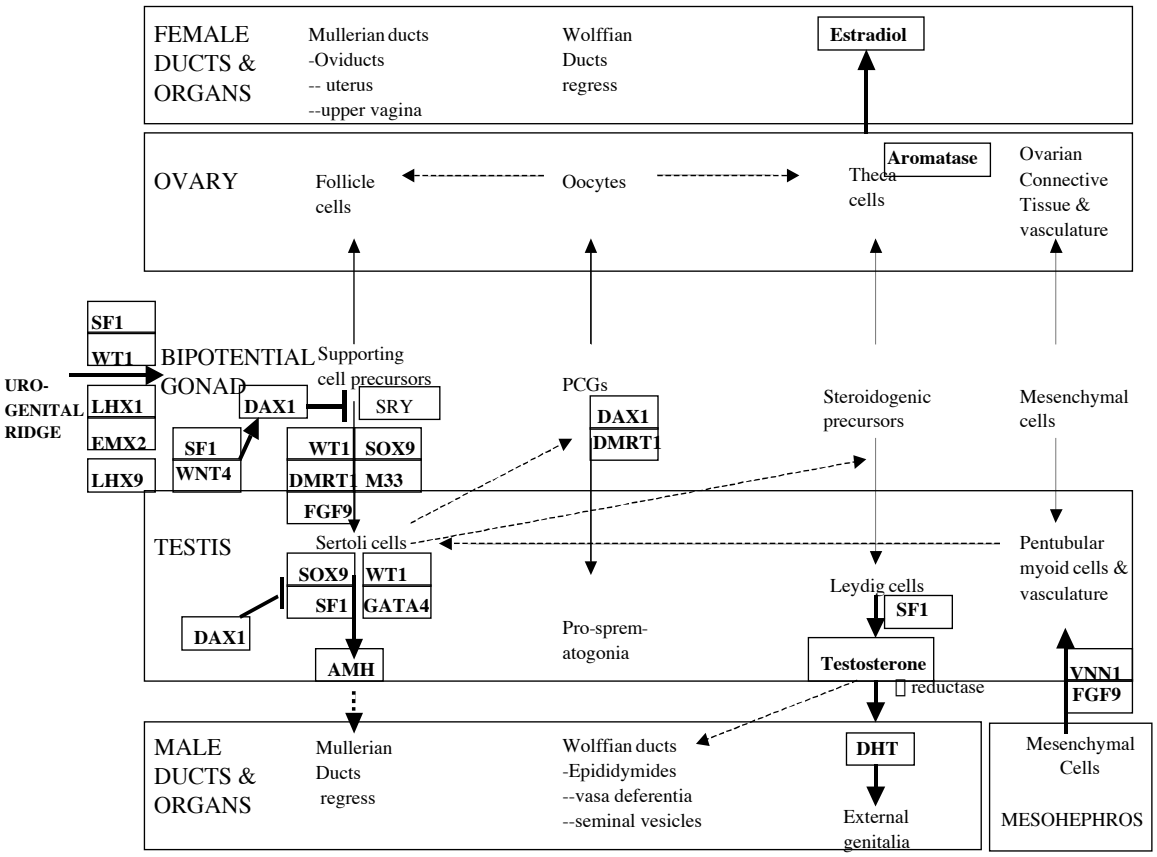
factor (*Tdy*) led researchers to a 35kb region of the Y chromosome now known to be the minimum necessary for male development. An intense research study culminated in 1990 in the identification of the sex-determining region (*Sry*) of the Y chromosome, found to be expressed in the genital ridge cells of the male mouse during the period in which testes begin to form [23]. The identity of *Sry* was confirmed by transgenic studies [24], in which female mice with two X chromosomes were made transgenic for the *Sry* gene, and developed as normal but sterile males with testes. Sry-like genes have been isolated from the males of a number of mammals [25] including rat, human, cattle [26] and voles [27]. However, in some lower vertebrates (birds, fish and amphibians) different mechanisms exist for sex determination. The avian females are heterogametic (ZW) and males homogametic (ZZ), which is the reverse of the situation found in mammals. And in alligators, temperature plays a role in sex determination [28].

Studies on the mouse XY gonad have found a marked increase in cell proliferation following the expression of the *Sry* gene at 11.25dpc, which is not mirrored in the XX gonad [29]. A reduced level of SRY in the developing mouse gonad, due to either a reduced number of SRY producing cells or a reduced level of SRY per cell, results in an ovarian pathway of gonad development [30]. The SRY protein is thought to act within the context of other gene products required for gonad development, and may act on or through one or more genes to ensure the differentiation and maintenance of Sertoli cells [31].

The human Y chromosome has been found to have approximately 50 genes (compared to approximately 1500 genes on the X chromosome), half of which are associated with male sex development and spermatogenesis [32]. Incidences of human

sex reversal (some familial) have lead investigators to try to understand the nature of the genetic abnormalities involved. Many of these sex reversal cases show mutations within the HMG box of the *Sry* gene (XY females) or the translocation of *Sry* to an X chromosome (XX males). However, 10% [33] of XX males do not have the *Sry* gene, and 75% of XY females have no detected *Sry* mutation. Many of these cases have been explained by the observation of mutations in other genes, thereby implicating them in sex reversal. For example, mutations in a gene thought to act downstream of *Sry, Sox9*, causes a condition known as Campomelic dysplasia, a skeletal malformation syndrome. Two thirds of XY individuals with sex reversal show Campomelic dysplasia [34]. Conversely, using a knock-in approach, the gene *Sox9* has been shown to substitute for *Sry* at the stage of gonad development [35]. From these and other studies, a picture is emerging of the possible pathways surrounding the action of the *Sry* gene.

Figure 1: The Jigsaw Puzzle of Vertebrate Sex Determination



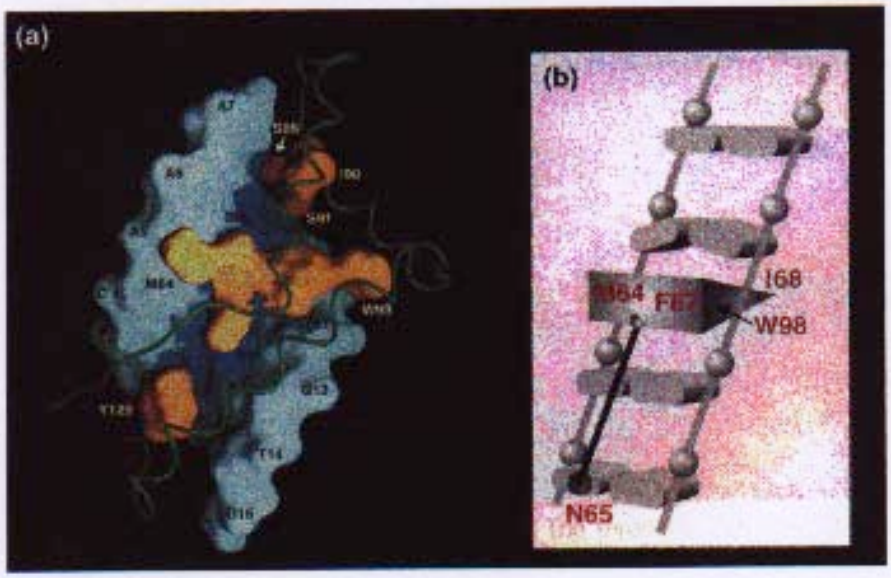
Vertebrate sex determination results from a complex network of regulatory interactions. The same basic set of genes appears to operate during early gonadal development in all vertebrate classes, despite the difference in mechanisms; and vertebrate sex determination results from a complex network of regulatory interactions. [1]

***Sox* Family of Genes**

Other members of the *Sox* gene family were identified through homology to a 79 amino acid motif known as the High Mobility Group (HMG) present in *Sry* [23] [36]. An HMG box of the *Sox* genes encodes an HMG domain with at least 50% amino acid identity with that of SRY. The SOX family are a subgroup of the HMG box proteins that show a highly specific tissue distribution and bind to identifiable DNA sequences with high affinities [37] [38]. However, this specificity appears to be context-dependent, that is, they appear to act in conjunction with other proteins [39] [40].

Figure 5 at the end of this section illustrates the homology of 19 of the mouse Sox gene HMG domains at the amino acid level. Classification of the HMG box proteins breaks them into two main families: (i) the MATA/TCF/SOX family and (ii) the UBF/HMG family. This partition of the HMG superfamily is directly related to the number of HMG boxes. Where there is more than one DNA binding domain, the binding occurs preferentially to bent DNA. This group include the HMG1 and HMG2 proteins. In the case of the *Sox* family, only one HMG domain is found. They may bind to pre-structured DNA with little or no specificity, but they also bind with high affinities in a sequence-specific manner to linear DNA [41]. The SOX HMG box binds specifically with the heptamer motif A/T A/T CAAAG of DNA, contacting the adenosines on both strands in the minor groove, causing the DNA helix to bend and opening up the helix [42]. As illustrated in Figure 2.

Figure 2: Protein-DNA interactions in the SRY-HMG/DNA



(a) Residues that directly interact with DNA bases are depicted as a yellow molecular surface attached to a green protein backbone. The intercalative wedge is formed by Met64, Asn65, Phe67, Ile68 and Try98, and intercalates between A-T basepairs 5 and 6. (b) Schematic representation of the wedge indicating how it inserts into the DNA. The wedge is anchored by Asn65, which hydrogen-bonds to the C4-G13 basepair below the site of insertion.

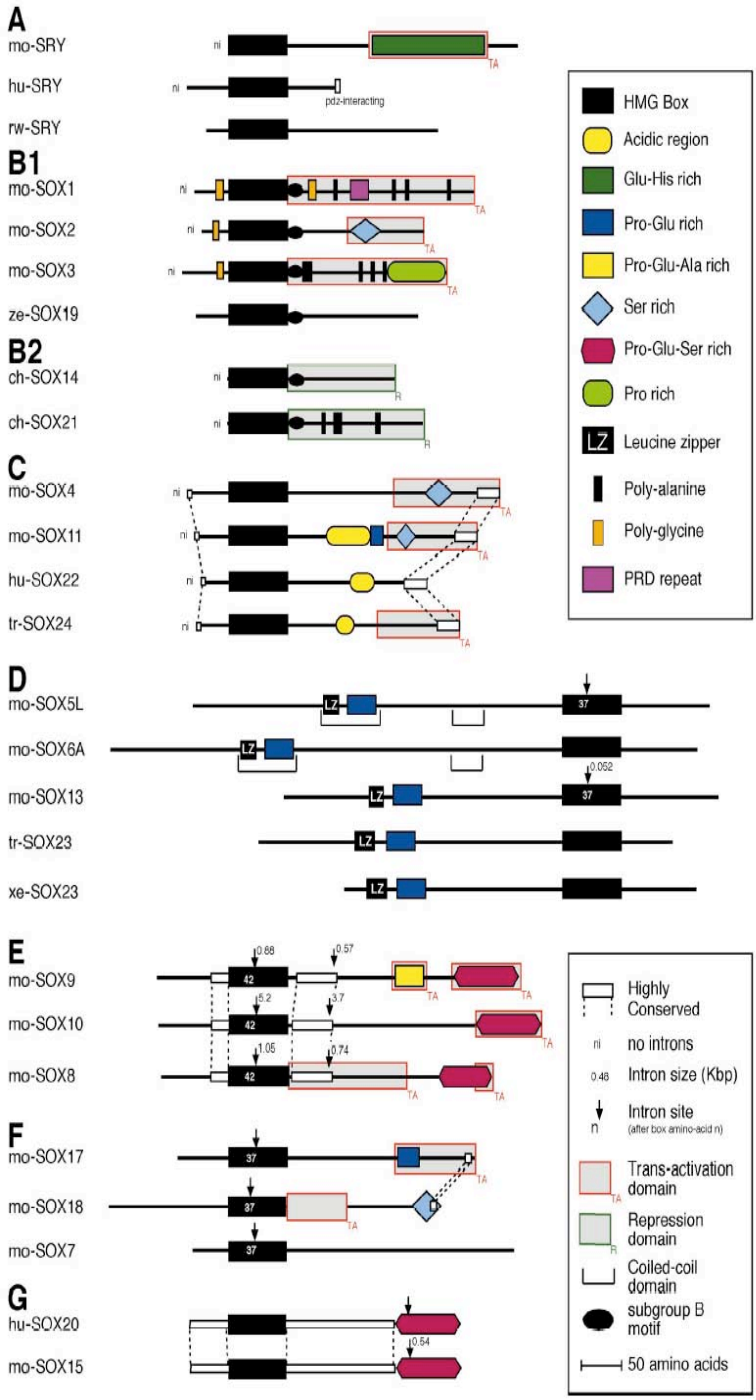
Ref:[43]

This altering of the local chromatin structure at specific sites may act to facilitate the interaction of distant enhancer nucleoprotein complexes with the transcription machinery [44] or the bending may serve to prevent the binding of such factors to adjacent sites in the major groove. This change in the architecture of the DNA is known to affect transcriptional elements vital to sex differentiation [45].

In addition to the HMG box, there are three other defining domains of *Sox* genes: a transactivating (TA) domain at the carboxy terminal, and distinctive C and N terminal domains [46]. A comparison of SOX protein domains and their groupings is illustrated below, together with the proposed phylogenic tree [2]. For some SOX groups, conservation of sequence and structure correlates with similarity of function, notably the B1 grouping which are all involved in the CNS (Central Nervous System). The phylogenic tree demonstrates the relatedness of the family of *Sox* genes between species, but it remains to be seen whether this relatedness is conserved at the molecular level in a mechanistic fashion.

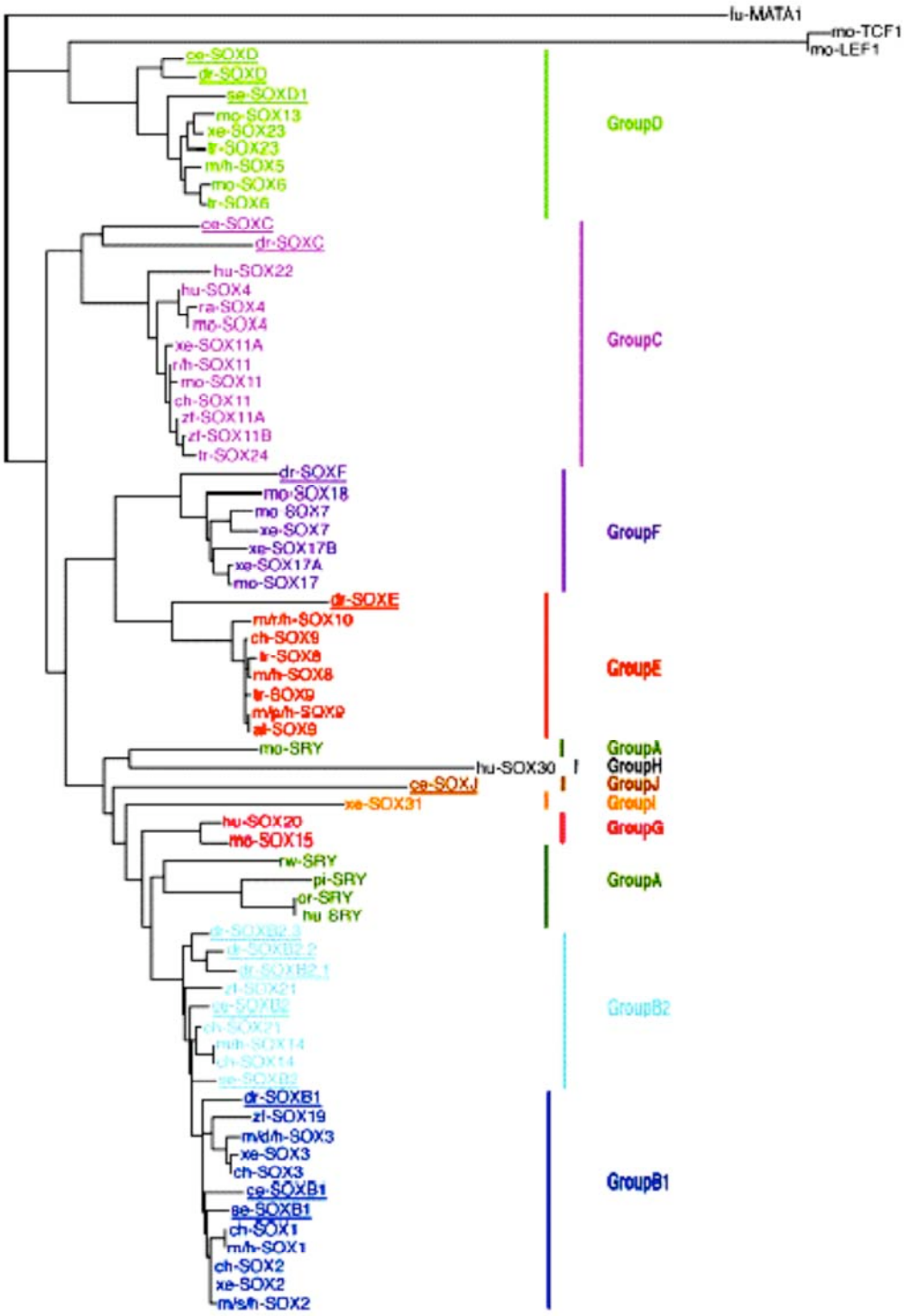
Through profiling the *Sox* genes across the mouse foetal collection of cDNAs in this project, a more detailed knowledge of expression will be recorded to add to the growing data on this interesting family of genes.

Figure 3: Schematic representation of SOX proteins



Diagram, highlighting conservation within SOX family groups. Proteins are arranged in groups as defined by HMG domain sequences. Various structural features, motifs, (demonstrated or putative) are shown along with intron positions and sizes where known. Genomic structures are known in some cases - 'ni' (no intron) indicates than an intronless structure has been reported[2].

Figure 4: Proposed Phylogeny of the SOX Transcription Factors



Tree created using the distance method FITCH (GCG), branch lengths are proportional to evolutionary distance or extent of divergence [2]

Figure 5: Sequence alignment of the HMG domains of the family of *Sox* genes.

```
SX12_MOUSE  .......... .......... .......MVW SQHERRKIMD QWPDMHNAEI SKRLGRRWQL LQDSEKIPFE REAERLRLKH MAD....... ..........
SX19_MOUSE  .......... .......... .......MVW SQIERRKIME QWPDMHNAEI SKRLGKRWKL LPDYEKIPFI KEAERLRLKH MA........ ..........
SOX9_MOUSE  .......... .......... .......MVW AQAARRKLAD QYPHLHNAEL SKTLGKLWRL LNESEKRPFV EEAERLRVQH KKD....... ..........
SX16_MOUSE  .......... .......... .......MVW SSAQRRQMAQ QNPKMHNSEI SKRLGAQWKL LDDEEKRPFV EEAKRLRARH LHDY...... ..........
SOX1_MOUSE  .........G GGTKANQDRV KRPMNAFMVW SRGQRRKMAQ ENPKMHNSEI SKRLGAEWKV MSEAEKRPFI DEAKRLRALH MKEHPDYKYR PRRK......
SOX2_MOUSE  .........G GNQKNSPDRV KRPMNAFMVW SRGQRRKMAQ ENPKMHNSEI SKRLGAEWKL LSETEKRPFI DEAKRLRALH MKEHPDYKYR PRRK......
SOX3_MOUSE  .........G GGGGSDQDRV KRPMNAFMVW SRGQRRKMAL ENPKMHNSEI SKRLGADWKL LTDAEKRPFI DEAKRLRAVH MKEYPDYKYR PRRK......
SX14_MOUSE  .......... ...SKPSDHI KRPMNAFMVW SRGQRRKMAQ ENPKMHNSEI SKRLGAEWKL LSEAEKRPYI DEAKRLRAQH MKEHPDYKYR PRRK......
SX15_MOUSE  .........G ASGGLPLEKV KRPMNAFMVW SSVQRRQMAQ QNPKMHNSEI SKRLGAQWKL LGDEEKRPFV EEAKRLRARH LRDYPDYKYR PRRK......
SOX8_MOUSE  .........G GGTLKAKPHV KRPMNAFMVW AQAARRKLAD QYPHLHNAEL SKTLGKLWRL LSESEKRPFV EEAERLRVQH KKDHPDYKYQ PRRR......
SX10_MOUSE  .........G AS..KSKPHV KRPMNAFMVW AQAARRKLAD QYPHLHNAEL SKTLGKLWRL LNESDKRPFI EEAERLRMQH KKDHPDYKYQ PRRR......
SOX4_MOUSE  .......... ........HI KRPMNAFMVW SQIERRKIME QSPDMHNAEI SKRLGKRWKL LKDSDKIPFI QEAERLRLKH MADYPDYKYR PRKK......
SX11_MOUSE  .......... ........HI KRPMNAFMVW SKIERRKIME QSPDMHNAEI SKRLGKRWKM LKDSEKIPFI REAERLRLKH MADYPDYKYR PRKK......
SOX7_MOUSE  .......... ...KSSESRI RRPMNAFMVW AKDERKRLAV QNPDLHNAEL SKMLGKSWKA LTLSQKRPYV DEAERLRLQH MQDYPNYKYR PRRK......
SX17_MOUSE  .......... ...AKAESRI RRPMNAFMVW AKDERKRLAQ QNPDLHNAEL SKMLGKSWKA LTLAEKRPFV EEAERLRVQH MQDHPNYKYR PRRR......
SX18_MOUSE  .........G ERQTADELRI RRPMNAFMVW AKDERKRLAQ QNPDLHNAVL SKMLGKAWKE LNTAEKRPFV EEAERLRVQH LRDHPNYKYR PRRK......
SOX5_MOUSE  VSESRIYRES RGRGSNEPHI KRPMNAFMVW AKDERRKILQ AFPDMHNSNI SKILGSRWKA MTNLEKQPYY EEQARLSKQH LEKYPDYKYK PRPKRTCLVD
SOX6_MOUSE  VAEARVYRDA RGRASSEPHI KRPMNAFMVW AKDERRKILQ AFPDMHNSNI SKILGSRWKS MSNQEKQPYY EEQARLSKIH LEKYPNYKYK PRPKRTCIVD
SX13_MOUSE  ....RHFSES R....NSSHI KRPMNAFMVW AKDERRKILQ AFPDMHNSSI SKILGSRWKS MTNQEKQPYY EEQARLSRQH LEKYPDYKYK PRPKRTCVVE
```

Sequence alignment of the HMG domain at the amino acid level for 19 members of the Sox family.

Objectives of my Research

To:

- create a panel of cDNAs from pure and intact RNA isolated from mouse foetal and placental tissues, to provide a reliable representation of the mouse tissue/organ transcriptome during development over the second half of gestation up to birth.
- demonstrate the validity and standardise the content of the panel.
- explore the specificity of the panel by PCR with primers designed to genes with known expression patterns in the adult mouse panel [20].
- illustrate the potential usefulness of the panel, using the newly emerging family of Sox genes as a model test system.
- compare the expression of two of the Sox gene members by PCR and *in-situ* hybridisation.
- review the process critically and examine alternative methods to explore gene expression during fetal development

References

http://www.ambion.com/techlib/basics/northerns/index.htm

1. Scherer, G., *The molecular genetic jigsaw puzzle of vertebrate sex determination and its missing pieces.* Novartis Found Symp 2002, 2002. **244**: p. 225- 236.
2. Bowles, J., G. Schepers, and P. Koopman, *Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators.* Developmental Biology, 2000. **227**(2): p. 239-255.
3. Gregory, S.G., et al., *A physical map of the mouse genome*. Nature, 2002. **418**(6899): p. 743-U3.
4. Powell, D., *Human Genome - Wellcome To The Genomic Age*. 2003, The Sanger Institute.
5. Waterston, R.H., et al., *Initial sequencing and comparative analysis of the mouse genome.* Nature, 2002. **420**(6915): p. 520-562.
6. Sudbeck, P., et al., *Sex reversal by loss of the C-terminal transactivation domain of human SOX9*. Nature Genetics, 1996: p. 230-232.
7. Kanai, Y., et al., *Identification of two Sox17 messenger RNA isoforms, with and without the high mobility group box region, and their differential expression in mouse spermatogenesis.* Journal of Cell Biology, 1996. **133**(3): p. 667-681.
8. Fuchs, P., et al., *Unusual 5' transcript complexity of plectin isoforms: novel tissue- specific exons modulate actin binding activity.* Human Molecular Genetics, 1999. **8**(13): p. 2461-2472.
9. Hastie, N., *Life, Sex, and WT1 Isoforms—Three Amino Acids Can Make All the Difference.* Cell, 2001. **106**: p. 391-394.

10. Mouchel, N., F. Broackes-Carter, and A. Harris, *Alternative 5' exons of the CFTR gene show developmental regulation.* Human Molecular Genetics, 2003. **12**(7): p. 759-769.

11. Kaufman, M.H., *The Atlas of Mouse Development*. 1998: Academic Press.

12. Docherty, K., *Gene Expression RNA Analysis*. Essential Techniques, ed. D. Rickwood. 1996: Wiley.

13. Bradley, A., *Mining the mouse genome - We have the draft sequence - but how do we unlock its secrets?* Nature, 2002. **420**(6915): p. 512-514.

14. Boguski, M.S., *Comparative genomics: The mouse that roared.* Nature, 2002. **420**(6915): p. 515-516.

15. Holloway, A.J., et al., *Options available - from start to finish - for obtaining data from DNA microarrays.* Nature Genetics, 2002. **32**: p. 481-489.

16. Smith, L., et al., *Single primer amplification (SPA) of cDNA for microarray expression analysis.* Nucleic Acids Research, 2003. **31**(3): p. art. no.-e9.

17. Miki, R., et al., *Delineating developmental and metabolic pathways in vivo by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays.* PNAS, 2001. **98**(5): p. 2199-2204.

18. Gerhold, D.L., R.V. Jensen, and S.R. Gullans, *Better therapeutics through microarrays.* Nature Genetics, 2002. **32**: p. 547-552.

19. Stoeckert, C.J., H.C. Causton, and C.A. Ball, *Microarray databases: standards and ontologies.* Nature Genetics, 2002. **32**: p. 469-473.

20. Freeman, T.C., et al., *Expression Mapping of Mouse Genes.* MGI Direct Data Submission, 1998.

21. Bush, T.G., et al., *Fulminant jejuno-ileitis following ablation of enteric glia in adult transgenic mice.* Cell, 1998. **93**(2): p. 189-201.

22. Takada, S., et al., *Delta-like and Gtl2 are reciprocally expressed, differentially methylated linked imprinted genes on mouse chromosome 12.* Current Biology, 2000. **10**(18): p. 1135-1138.

23. Gubbay, J., et al., *A gene-mapping to the sex-determining region of the mouse y-chromosome is a member of a novel family of embryonically expressed genes.* Nature, 1990: p. 245-250.

24. Koopman, P., et al., *Male Development of Chromosomally Female Mice Transgenic for Sry.* Nature, 1991. **351**(6322): p. 117-121.

25. Margarit, E., et al., *Identification of conserved potentially-regulatory sequences of the SRY gene from 10 different species of mammals.* Biochem. Biophys. Res. Commun., 1998. **245**(2): p. 370 - 377.

26. Liu, W., et al., *A radiation hybrid map for the bovine Y Chromosome.* Mamm Genome, 2002. **13**(6): p. 320-6.

27. Fernandez, R., et al., *Mapping the SRY in Microtus cabrerae: a vole species with multiple SRY copies in males and females.* Genome, 2002. **45**(2): p. 600 - 603.

28. Morrish, B. and A. Sinclair, *Vertebrate sex determination; many means to an end.* Reproduction, 2002. **124**(4): p. 447- 457.

29. Schmahl, J., et al., *Sry induces cell proliferation in the mouse gonad.* Development, 1999. **127**: p. 65 - 73.

30. Washburn, L.L., K.H. Albrecht, and E.M. Eicher, *C57BL/6J-T-Associated Sex Reversal in Mice Is Caused by Reduced Expression of a Mus domesticus Sry Allele.* Genetics, 2001. **158**: p. 1675 - 1681.

31. Lovell-Badge, R., C. Canning, and R. Sekido, *Sex-determining genes in mice: building pathways.* Genetics and Biology of Sex Determination, 2002. **244**: p. 4-22.

32. Graves, J.A.M., *The rise and fall of SRY.* Trends in Genetics, 2002. **18**(5): p. 259 - 264.

33. Boucekkine, C., et al., *Clinical and Anatomical Spectrum in Xx Sex Reversed Patients - Relationship to the Presence of Y-Specific DNA-Sequences.* Clinical Endocrinology, 1994. **40**(6): p. 733-742.

34. Meyer, J., et al., *Mutational analysis of the SOX9 gene in campomelic dysplasia and autosomal sex reversal: Lack of genotype/phenotype correlations.* Human Molecular Genetics, 1997: p. 91-98.

35. Videl, V.P.I., et al., *Sox9 induces testis development in XX transgenic mice.* Nature, 2001. **28**(3): p. 216-217.

36. Denny, P., et al., *A conserved family of genes related to the testis determining gene, SRY.* Nucleic Acids Res, 1992. **20**(11): p. 2887.

37. Harley, V.R., et al., *The HMG box of SRY is a calmodulin binding domain.* Febs Letters, 1996. **391**(1-2): p. 24-28.

38. Pevny, L.H. and R. LovellBadge, *Sox genes find their feet.* Current Opinion in Genetics & Development, 1997. **7**(3): p. 338-344.

39. Kamachi, Y., M. Uchikawa, and H. Kondoh, *Pairing SOX off with partners in the regulation of embryonic development.* TRENDS IN GENETICS Osaka 5650871, Japan Osaka Univ, Inst Mol & Cellular Biol, Suita, Osaka 5650871, Japan, 2000. **16**(4): p. 182-187.

40. Wilson, M. and P. Koopman, *Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators*. Current Opinion in Genetics & Development, 2002. **12**(4): p. 441-446.

41. Soullier, S., et al., *Diversification pattern of the HMG and SOX family members during evolution.* Journal of Molecular Evolution, 1999. **48**(5): p. 517-527.

42. van de Wetering, M. and H. Clevers, *Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix.* EMBO J, 1992. **11**(8): p. 3039-44.

43. Werner MH, et al., *Molecular determinants of mammmalian sex*. Trends in Biochem Sci, 1996. **8**: p. 302 - 208.

44. Vaccari, T., et al., *Hmg4, a new member of the Hmg1/2 gene family.* Genomics, 1998: p. 247-252.

45. Harley, V.R., et al., *DNA binding activity of recombinant SRY from normal males and XY females.* Science, 1992. **255**(5043): p. 453-6.

46. Kamachi, Y., K.S.E. Cheah, and H. Kondoh, *Mechanism of regulatory target selection by the SOX high-mobility- group domain proteins as revealed by comparison of SOX1/2/3 and SOX9.* Molecular and Cellular Biology, 1999. **19**(1): p. 107-120.

CHAPTER 1 Construction of the Panel

Introduction

Mouse material was obtained from the Anatomy Department, Cambridge University and brought to the Sanger Centre on cardice for subsequent RNA extraction, reverse transcription, panel construction and analysis by PCR.

1.1. Special Equipment and Suppliers

Dissecting kit	
Dissecting microscope	
Powergen 700 homogeniser	Fisher Scientific
Metal block and hammer	made in-house
Multichannel pipettors	Matrix range
RNA- and DNA-free eppendorf tubes and pestles	CP labs
15 ml falcon tube (individually wrapped)	Falcon # 2059
UV Spectrophotometer	Amersham Pharmacia
Flatbed electrophoresis equipment	Flowgen and Hybaid
UV transilluminator	Genetic Research Instrumentation
Polaroid camera	Genetic Research Instrumentation
Film for Polaroid camera	Genetic Research Instrumentation
Image capture software	developed by Alan Flook
Microtitre plates	CostarCorning # 6511
Hybaid rubber mats	Hybaid
MJ tetrad thermocyclers	Genetic Research Instrumentation

1.2 Materials and Solutions

Chemicals and Suppliers

All reagents are either analar or molecular biology grade and made up in double deionised water at 18 ohms, unless otherwise stated.

1 kb DNA ladder	GibcoBRL # 15615-024
8-hydroxyquinoline	Sigma # H6878
agarose	GibcoBRL # 15510-019
amplitaq	Perkin Elmer #
boric acid	Sigma # B 7660
bovine serum albumin fraction V	Sigma # A8022
bromophenol blue	Sigma # B5525
chloroform	Sigma # C 5312
citric acid	Sigma # C 3674
Cresol Red	ALDRICH # 11448-0
deoxyribonuclease 1 amplification grade (Dnase 1)	GibcoBRL # 18068-015
diethyl pyrocarbonate (DEPC)	Sigma # D 5758
dithiothreitol	Sigma # D9779
ethanol	BDH # 10107
ethidium bromide solution 10 mg/ml (EtBr)	Sigma # E1510
ethylene diamine tetraacetic acid (EDTA)	Sigma # ED2SS
glycerol	Sigma # G 7757
human placenta DNA	Sigma # D3160
isopentane	Sigma # 27,041-5
mixed bed resin TMD-8	Sigma # M-8157

moloney murine leukemia virus	
reverse transcriptase (M-MLV)	GibcoBRL # 28025-021
mouse gDNA	Clontech # 6650-1
oyster glycogen	Sigma # G 8751
phenol ultra pure	GibcoBRL # 5509UB
phosphate buffered saline (PBS)	Oxoid #BR14
primers	Sanger Institute
rat gDNA	Promega # G313A
RNAce total pure purification system	Bioline # BIO-28054
Rnase Inhibitor (RNAguard)	Pharmacia # 27-0815-01
sodium acetate	Sigma # S 2889
"T17 primers (T17A,G,C)"	Sanger Institute
trizma base	Sigma # T 1503
ultrapure dNTP 2'-deoxyribonucleoside 5'-triphosphate	
100 mM	Pharmacia # 27-2035-
01xylene cyanol FF	Sigma # X 4126

1.2.1. depc-treated water

Diethylpyrocarbobate (depc) is an alkylating agent and inactivates any proteins including RNAses found in solutions. It is a potential carcinogen and all manipulations should be performed in a fume hood.

- Add 1.8 ml depc to 1.8 L of nanopure (18 ohm) millipore water in a 2 L duran bottle.

- Shake vigorously, cover the lids with foil, autoclave tape and label allow to stand overnight in the fume-hood.
- Autoclave. Upon autoclaving, depc is broken down to carbon dioxide and ethanol.

1.2.2. 1 kb ladder

1 ml 1 kb DNA ladder (GibcoBRL)

7.6 ml depc treated water

1.68 ml 6x LB

Dispensed as 400 µl aliquots, store at –20°C

Fragments sizes (bp): 12,216, 11,198, 10,180, 9,162, 8,144, 7,125, 6,108, 5,090, 4,072, 3,054, 2,036, 1,635, 1,018, 516/506, 394, 344, 298, 220, 200, 154, 142, 75

1.2.3. 6 x LB	fc (final concentration)
6 ml glycerol	33%
14 ml depc treated water	66%
0.05 g bromophenol blue	0.25%

filtered through 0.22um filter,

Store frozen as 2 ml aliquots

1.2.4. T0.1E	fc
121.1 g Trizma base	1 M
37.22 g EDTA	0.1 M

pH 8 with hydrochloric acid and made up to 1 L with depc treated water

1.2.5. TBE 10 x Stock fc

108 g Trizma base	890 mM
55 g Boric acid	890 mM
40 ml 0.5M EDTA	20 mM

pH 8 with hydrochloric acid and made up to 1 L with depc treated water

1.2.6. Cresol Red Solution fc

0.845 mg/ml Cresol Red sodium salts in T0.1E	2mM

1.2.7. 10 x PCR Rxn buffer (3.5 mM $MgCl_2$fc) fc 50 x 2 ml

4.50 mls 1 M Tris-HCl pH 8.8	100 mM	45 ml
5.00 mls Cresol Red solution	0. 1mM	50 ml
0.15 ml depc treated water		1.5 ml
0.35 ml 1 M $MgCl_2$	3.5 mM	3.5 ml
0.1454 g $(HN_4)_2SO_4$	10 mM	1.454

Store frozen in 2 ml aliquots

1.2.8. Dilution buffer 48 x 3.1 ml

16.00 mls depc treated water	100 ml
8.00 mls T0.1E	50 ml
0.13 ml Cresol red solution	0.8125 ml
14 l 1 M NaOH (to pH 8.1)	0.0875 ml

Store frozen in 3.1ml aliquots

1.2.9. Sucrose Solution fc

121.1 g sucrose in 350 ml depc treated water 34.6%

Store frozen in 7ml aliquots

1.2.10. PCR Standards: fc

Blank: oyster glycogen (Sigma) 2 □g/□l

Genomic mouse DNA 4 ng/□l

Genomic human placenta DNA 4 ng/□l

Genomic rat DNA 4 ng/□l

1.2.11. 10 mM Deoxynucleotides fc

Ultra pure dNTP set PHARMACIA # 27-2035-01

4 x 250 □l 100 mM 2'-deoxyadenosine 5'-triphospate 10 mM

4 x 250 □l 100 mM 2'-deoxycytidine 5'-triphospate 10 mM

4 x 250 □l 100 mM 2'-deoxyguanosine 5'-triphospate 10 mM

4 x 250 □l 100 mM 2'-deoxythymidine 5'-triphospate 10 mM

6 ml depc treated water

Store frozen in 1ml aliquots

1.2.12. Phosphate Buffered Saline (PBS)

1 tablet (Oxoid Dulbecco A) to 100 ml double distilled water

autoclaved, 115^0C, 10 mins

1.3. Methods

1.3.1. Preparation and maintenance of Animals

BLJ6 females and stud males, obtained from Harlan, were maintained in a twelve hour light/dark cycle with controlled temperature and allowed free access to food and water, and were naturally mated in the Combined Biofacility of the School of Biological Sciences, Cambridge University.

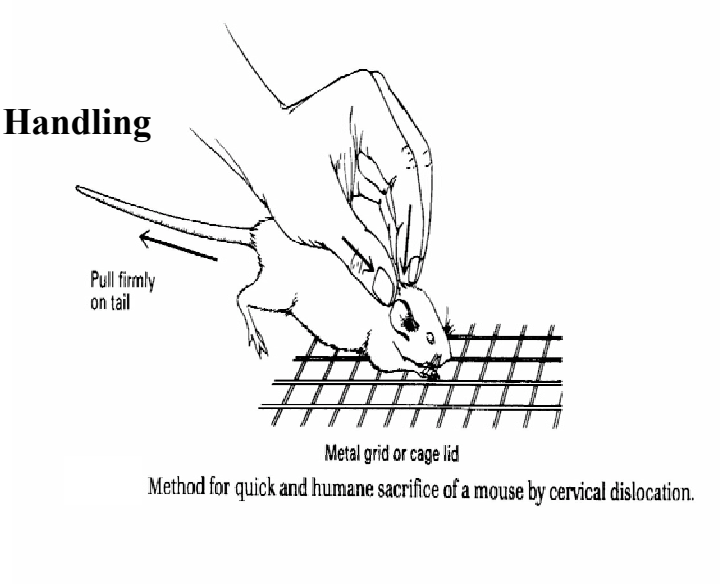
1.3.2. Tissues Collection

Mouse embryo dissections were performed by Professor Martin Johnson and Dr Anne Ferguson-Smith. Dissections were carried out over a period of 12 months to collect sufficient material for the panel, using gloved hands and thoroughly cleaned dissecting instruments. All mice were culled between 20.00 and 22.00 hours.

In brief:

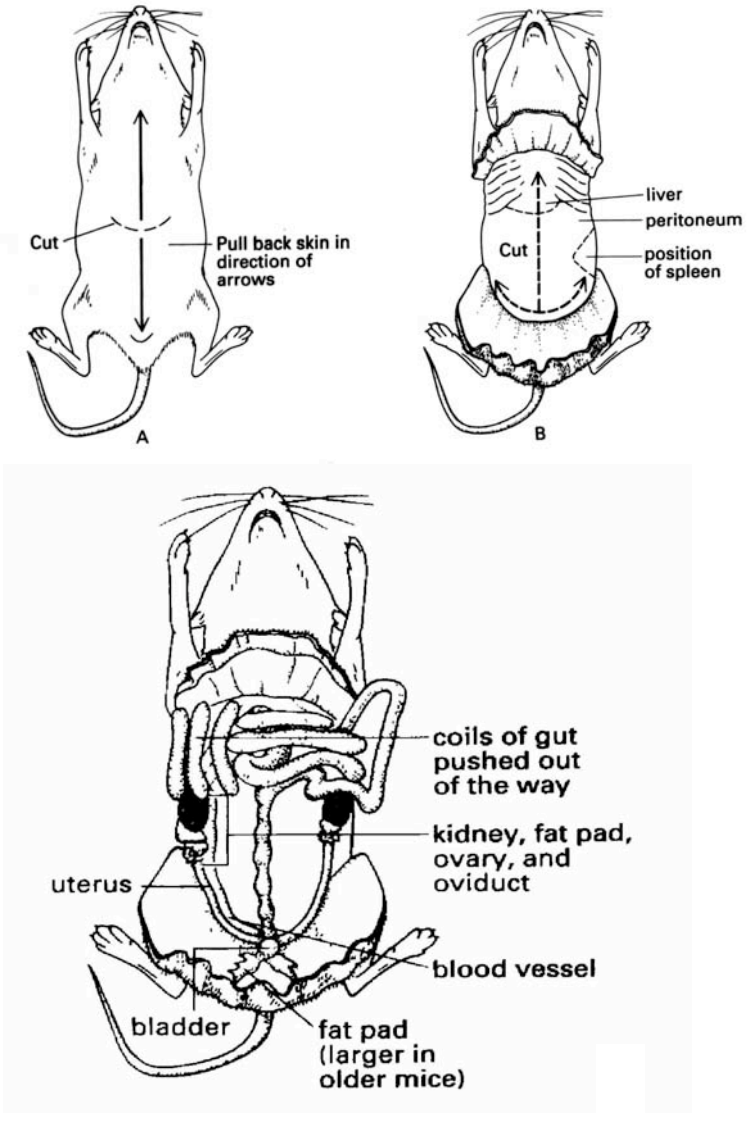
- Females naturally mated with male mice were identified by the presence of a vaginal plug (= d0.5), and killed by cervical dislocation (see Figures 6) at the following stages: 8.5d, 9.5d, 10.5d, 11.5d, 12.5d, 13.5d, 15.5d, 17.5 days post-coitum (dpc), and their uterine horns excised for dissection (see Figure 7).
- The foetuses were removed into sterile phosphate buffered saline (PBS) at room temperature. And photographed (see appendix 1).
- They were then dissected under the microscope, using flame polished forceps into cold sterile PBS.
- Tissues were placed into isopentane held on cardice and finally stored in 0.5 ml sterile tubes in liquid nitrogen prior to RNA extraction.

Figure 6 : Mouse Handling



Method for quick and humane sacrifice of a mouse by cervical dislocation.

Figure 7: Dissection of Mice



1.3.3. Dissection Procedure

- Taking one foetus at a time, the outer uterine musculature was removed keeping the placenta intact and the amnion/yolk sac was peeled away before taking a photographic record of the developmental stage.
- The placenta/yolk sac was separated from the foetus proper, cutting away the umbilicus leaving the fetal gut intact; the placenta and yolk sac were then removed separately.
- Cutting near the shoulder and hip joints the limbs were removed.
- The abdomen was then cut open to remove the liver, the duodenum, then the stomach, spleen and lastly separating the intestines. The mesentery was removed in all cases.
- The heart was removed from the thoracic cavity, followed by the release and separation of the oesophagus and lungs.
- Gonads were then removed, together with the urogenital tract (mesonephros) and where possible, in later stages, differentiated into male and female gonads.
- Next the bladder, adrenals, and kidneys were removed.
- The brain was exposed through making an incision at the back of the neck and gently peeling back the skin over the head, to remove the cranium. The whole brain and spinal cord, where then teased out and separated into the forebrain, midbrain, hindbrain and spinal cord.
- The rest of the head was then taken, and finally the remainder of the body was collected as a single sample.

1.3.4. Extraction of RNA from Tissues using Bioline RNace Kit

General notes: Skin surfaces and dust particles contain an abundance of RNAse, so it is important to wear gloves throughout and change them frequently, also to keep bottles/tip boxes/tubes covered as much as possible. Use RNAse free plastic and plugged tips, to prevent contamination. Chemicals used for RNA work should be kept separate from the general stock and weighed out without the use of spatulas, by carefully tapping the chemical out of the bottle.

Solutions and water should be treated with Diethyl PyroCarbonate (depc) see section 1.2.1. Tris solutions, however require making up in depc treated water, as they do not respond well to neat depc. Solutions, which cannot be autoclaved, can be filtered through a sterile 0.22 μm filter to reduce contaminating ribonucleases.

Equipment can be cleaned by soaking in 0.1 M NaOH + 1 mM EDTA for 1 hour followed by rinsing in depc treated water.

Small scale (5 – 50 mg)

- Place tissue in a sterile RNAse-free 1.5 ml eppendorf tube, add 350 μl of lysis solution homogenise using a sterile RNAse-free pestle.
- Add 25 μl of adsorbin solution (mix well before use) and vortex. Place on ice for 5 mins.
- Centrifuge for 2 mins to sediment the adsorbin pellet. Carefully remove the aqueous layer without disturbing the gelatinous pellet and place in a new sterile 2 ml eppendorf tube.
- Add 700 μl water saturated phenol plus hydroxyquinoline, then add 140 μl chloroform and 70 μl buffer A.
- Vortex for 10 secs place on ice for a further 5 mins.

- Spin 14,000 rpm for 10 mins at 4^{o}C, remove the upper aqueous layer being careful not to remove any of the interface. Place in a new sterile 1.5 ml eppendorf tube.
- Add an equal volume of isopropanol, mix and leave at –20^{o}C for a minimum of 20 mins.
- Spin at 14,000 rpm for 10 mins at 4^{o}C, remove the supernatant.
- Add 1 ml ice-cold 75 % ethanol, gently agitate through a tip to wash the pellet, spin at 14,000 rpm for 5 mins. Repeat twice, removing as much ethanol as possible on the last wash without dislodging the pellet.
- Allow the pellet to air dry for 15 mins, re-suspend in 50 µl depc treated water and measure the OD.
- Take 2 µl of the RNA solution to 1 ml T0.1E and measure at 260nm and 280nm. Take note of the 260/280 ratio and calculate the concentration of total RNA (1 OD at 260 nm is equivalent to 40 µg/ml, a ratio of 1.8-2 is preferred, though ratios as low as 1.5 are still worth pursuing).
- Add 150 µl 100 % ethanol to the original 50 µl RNA solution and make a new calculation of the concentration.
- Remove 2 µg to a sterile 0.5 ml eppendorf tube, add 1/10th the volume as 3 M sodium acetate, pH 5.2, leave at –20^{o}C for 30 mins. Pellet at 14,000 rpm for 5 mins remove the supernatant, air dry, add 7 µl of 1 X LB and run on a 1 % agarose + 0.004 % Ethidium Bromide gel, in 1 X TBE. Run at 30 milliamps, 65 volts for 1.5 to 2 hours in 1 X TBE + 0.004% Ethidium Bromide.

1.3.5. Deoxyribonuclease Treatment

Take 50 µg of total RNA from the ethanol stock add 1/10th the volume as 3 M sodium acetate, pH 5.2, leave at –20^{o}C for 30 mins. Pellet at 14,000 rpm for 5 mins, wash with 70 % ethanol, leave to air-dry.

		fc
50 µg	RNA pelletted, washed and dried	1 µg/µl
39.5 µl	depc treated water	
0.5 µl	RNA Guard 33 U/µl (Pharmacia)	0.3 U/µg
5 µl	10 X DNase I reaction buffer (BRL)	x 1
5 µl	Dnase 1 U/µl (BRL)	0.1 U/µg
50µl	Total	

leave at room temperature for 15 mins

add 5 µl 25 mM EDTA (BRL)

65^{o}C for 10 mins, 90^{o}C for 2mins, Chill on ice, spin

1.3.6. Reverse Transcription

		fc
50 µl	DNA-free RNA (50 µg see section 1.3.5.)	250 ng/µl
40 µl	5 x first strand buffer (BRL)	x1
5 µl	10 mM dNTPs	250 µM
20 µl	100 mM DTT (BRL)	10 mM
15 µl	200 U/µl M-MLV (BRL)	15 U/µl or 60 U/ng RNA
20 µl	40 mM Sodium Pyrophosphate	4 mM
50 µl	20 µM Primer T_{17}AGC	5 µM
200 µl		

37^{o}C for 90 mins, 60^{o}C for 5 mins, Chill on ice, spin

1.3.7. Stock and working solutions

cDNA currently at 250 ng/µl (50 µg/200 µl), was diluted (1/5) with 800 µl T0.1E to make a 1 ml stock solution of 50 ng/µl and stored as 50 µl aliquots at -20°C. 50 µl aliquots plus 1.2 ml (1/25) oyster glycogen solution at 2 mg/ml in depc treated water to provide a cDNA template working solution for PCR at 2 ng/µl.

1.3.8. Polymerase Chain Reaction

per reaction		fc
7.2 µl	34.6 % sucrose	12.45 %
0.187 µl	1/10 fresh β-mercaptoethanol in T0.1E	0.093 %
1.0 µl	10 mM dNTPs	0.5 mM
0.125 µl	AmpliTaq 5 U/µl	0.03125 U/µl
2.0 µl	10 x buffer 35 mM $MgCl_2$	3.5 mM
3.488 µl	Dilution buffer	0.228 µg/µl (cresol red)
1.0 µl	forward and reverse primers at 100 ng/µl	5 ng/µl
5.0 µl	RT product/cDNA template at 2 ng/µl	0.5 ng/µl
20 µl		

1.3.9. Typical Cycling Programme

Using hot lids

92°C 2 mins		
92°C 30 secs	denaturation)	
55°C 90 secs	annealing)	35 cycles
72°C 1 min	polymerasation)	
72°C 10 mins	extension	

1.3.10. Primer design

Primers were designed using the computer program PRIMER (version 0.5, Whitehead Institute for Biomedical Research, 1991) by Sarah Hunt, to regions 300 bp from the 3' end of the sequence, with a GC content of 40-60% and blast searched for complementarities to other known sequences. They were then synthesised by Geneset (Paris, France) or in-house at the Sanger Centre. Figure 8 identifies the primer sequences used in the purity testing of the panel.

Where primers were synthesised in-house at the Sanger Centre, they were supplied in ammonium hydroxide, dried down in a speedie vac and suspended in milliQ water to a concentration of 1 mg/ml as assessed by OD at 260 nm. Primer mixes of forward and reverse at 100 ng/μl were then made as working solutions and stored at –20^{o}C.

The planned dissection and tissue collection procedures were modified during the initial few dissections, ensuring that all dissectible tissues (defined as those tissues developed sufficiently to be unambiguously distinguished as those tissues) were collected in a given sequence.

The RNA extraction procedure was developed from Bioline's original protocol. RTPCR conditions were developed in-house originally for the adult mouse panel and subsequently used for a number of panels created in the Gene Expression Group. Similarly panel construction was a method devised in-house and applied to this foetal panel.

Figure 8: Primer Sequences

Sanger ID	Forward Primer	Reverse Primer	Symbol	Accession number
st95_22	ATGTGTGTGTGTGCACATG	TACACCCGCACTAATGGTCA	Slc17a2	L33878
st95_35	GCTTCTATCTGGCGGAAGG	TGTCATCTGGCTACCTTCCC	Calb1	M21531
st95_42	GCTTCTATCTGGCGGAAGG	TGTCATCTGGCTACCTTCCC	Calb1	M23663
st95_43	CTCTTTTCTCCCACCTCATCC	CTACCAGGCAGCAGGAGTTC	Hox5b	M26283
st95_58	AAGTCAACTTCTCAGAGCCTGG	GCTTTGACAAGGCTGGAGAC	Fsbpi	M65034
st95_69	CCCAGAAGTACTACGGGAAGG	CGAGTTGCCGTGTGTGAG	Adora1	U05671
st96_128	GCCAATGATTCATCTTGAGTTG	CCTTGATTCTCTCCGCTCAG	Csna	M36780
st96_153	CTCATGGTTTTCCCCTCTGA	GGTTCTGCTTTATTGACCTTGG	Fabph1	U02883
st96_273	GGGGAAGTGGAACACACGG	AGCAGGAGTTGGCTGGAATG	Si-s	X15546
st96_310	CTGATCCGCAAATACGGG	GCATGATCGGTTCCACTTG	Rps29	L31609
st96_428	GAAGAGTTCTGAGCATGCCC	TTCTTGGGGCCTATGGAAG	Cab45	U45978
st97_760	ATGGCTTGATTGGTACCAGTG	GACAAGTGGAAAAACAGGAAGC	bac 573K1a	
st97_761	TGCAGGCAGAGATGCTACTG	CGCTCAGAGAGAAAAAAATTGG	bac 573K1b	

1.4. Construction of the Panel

1.4.1. Tissue dissection

As mentioned in section 1.3.2., tissue was collected over a period of 12 months, stored frozen and transported to the Sanger Centre for RNA extractions. Figure 9 below illustrates the information collected at the time of tissue dissection.

Figure 9: Table of Tissues Dissected

Tissues	Plug date	Kill date	Kill time	Female no	Gestation/ Morphological age	Photo log no	sex	Tray code
Placenta								
Yolk sac*								
Liver								
Oesophagus/stomach*								
Intestines*								
Spleen*								
Bladder*								
Kidney*								
Female *Urogenital + gonad								
Male *Urogenital + gonad								
Heart*								
Lung*								
Forelimb								
Forebrain*								
Midbrain*								
Hindbrain*								
Spinal cord*								
Rest of head								
Remainder of body								

The asterisk, denote material combined from an entire pregnancy, therefore the amount of a given tissue was dependent on the foetal number in addition to the developmental stage. A total of 60 mice were culled providing 425 foetuses, with approximately 8000 dissected tissue samples. From this collection, RNA was

extracted from a single pregnancy at any one time, combining material where appropriate.

1.4.2. RNA extraction and assessment

A 50 μg, good quality RNA sample, as defined by ratio of ODs and strength of 18S/28S banding on an agarose gel, was selected. Figure 10 demonstrates a typical table showing this information for a 15.5 day foetus dissected on 19.11.97.

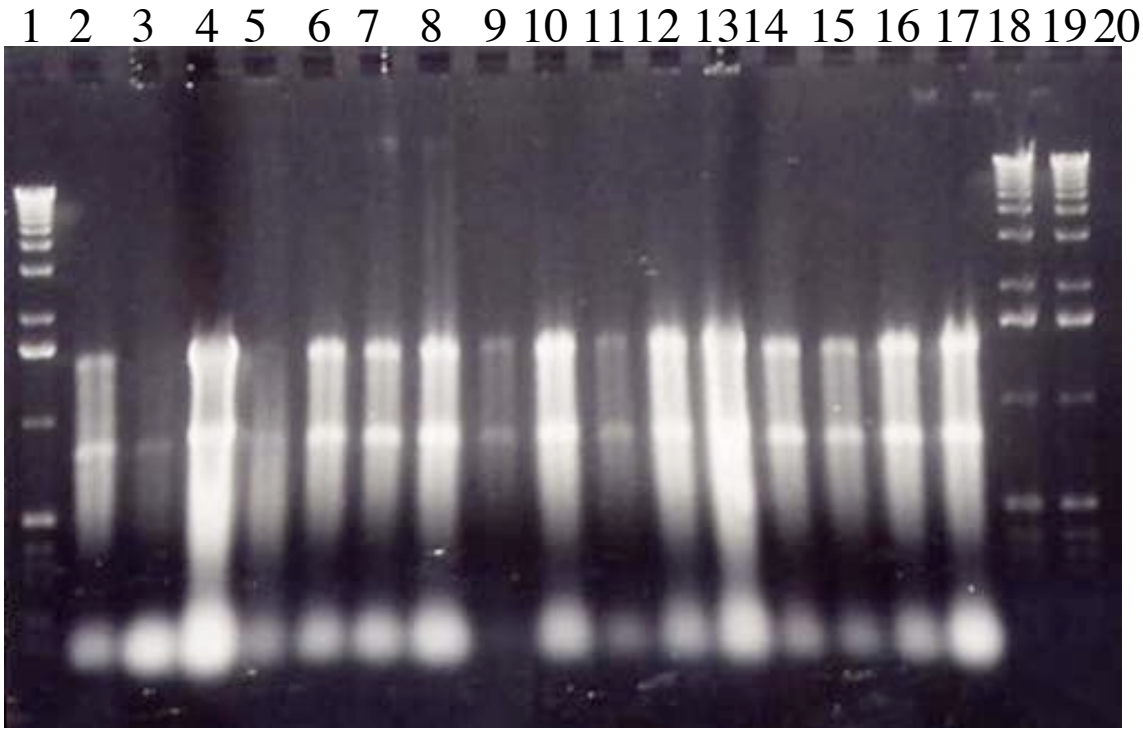
Figure 10: RNA extraction from 15.5d foetus

Tissues 15.5 day female date:19.11.97	Weight mg	Sample number	260nm	280nm	260/280 ratio	μg/μl	Total amount of RNA μg	ng RNA per mg tissue	Quality post agarose gel
Placenta	400	4	0.016	0.009	1.77	0.32	64	160	Good
Yolk sac	505	9	0.049	0.03	1.63	0.98	49	97	Poor
Liver	120	4	0.086	0.049	1.75	1.72	86	716	Good
Oesophagus/stomach	60	9	0.066	0.043	1.53	1.32	66	1100	Poor
Intestines	122	9	0.177	0.078	1.5	2.34	117	959	Good
Spleen	30	9	0.002	0.001	2	0.04	2	66	
Bladder	50	9	0.007	0.008	0.87	0.14	7	140	Good
Kidney	60	9	0.03	0.025	1.2	0.6	30	500	Good
Female Urogenital + gonad	10	1	0	0.005					
Male Urogenital + gonad									
Heart	60	9	0.347	0.232	1.48	6.9	345	5750	Good
Lung	210	9	0.217	0.115	1.88	4.34	217	1033	Good
Forelimb	90	1	0.018	0.031	0.88	0.36	18	200	Good
Forebrain	240	9	0.125	0.062	2.01	2.5	125	520	Good
Midbrain	170	9	0.047	0.033	1.42	0.94	47	276	Good
Hindbrain	160	9	0.032	0.021	1.52	0.64	32	200	Good
Spinal cord	50	9	0.025	0.016	1.56	0.5	25	500	Good
Rest of head	160	2	0.251	0.148	1.7	5.02	251	1568	Good
Remainder of body	400	2	0.065	0.038	1.71	1.3	65	162	Good

A typical gel of 2 μg amounts of RNA as assessed by OD measurements of the above tissue extraction is shown below as Figure 11. Samples are in the same order as the above table, excluding spleen and urogenital samples. From this gel, the quality of the RNA is scored and an indication as to the accuracy of the OD measurements gained. Repeat gels are run for confirmation, especially where the lanes are overloaded (lanes 3, 5,6,7,9,11,12, 15 and 16). The sample in lane 3 (yolk sac) is a

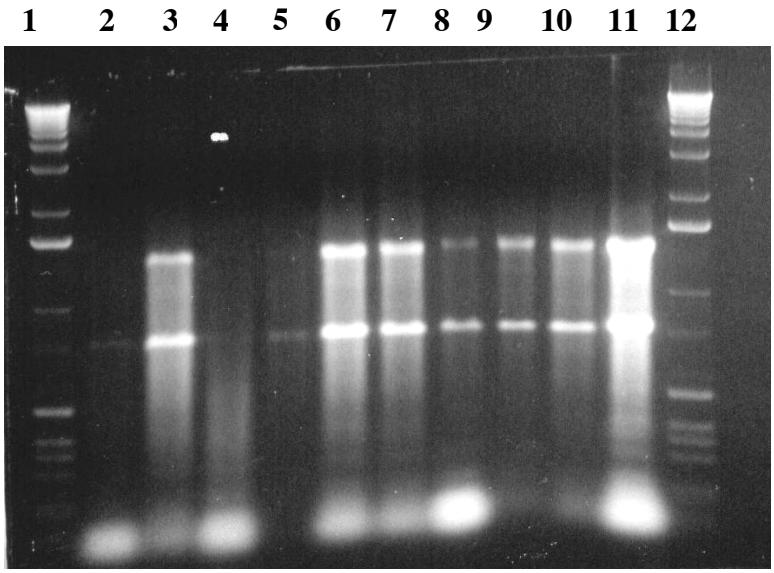
poor sample and was subsequently re-extracted. In this instance, the placental sample (first lane of gel) was given a score of 1, all other RNA volumes were then adjusted in relationship to this, to provide an approximately equivalent quantity of total RNA in each tissue sample.

Figure 11: Agarose gel of RNA extracted from 15.5d foetus.



A good quality RNA profile on a gel should show no contaminating DNA (near the gel origin), bold 18S and 28S bands with a stronger representation at the 28S level with minimal banding at the 5S level. The full range of tissues from a single gestation stage was extracted for total RNA, in this fashion on a single day. The process was repeated until there was sufficient material for the completed list of tissues and stages of animals [see above about numbers. of foetuses contributing to each sample]. Total RNA was stored at –70^{o}C in 75% ethanol as a suspension.

Figure 12: Illustration of differing qualities of RNA



Lanes 1 & 12: 1kb ladder

Lanes 2, 4 & 5: 18s band very faint, little RNA present - poor quality – not used.

Lane 8: 28s band much lighter than 18s, - degraded RNA at bottom of gel – not good, could be used if no other sample available.

Lanes 3, 6, 7 & 11: some degradation in lane 11, but generally strong 28s and 18s banding, good smearing between the two major bands where the message is expected to migrate – good quality.

Lanes 9 & 10: Clear 18s and 28s bands, no obvious degradation at bottom of gel – good quality.

1.4.3. Reverse transcription and PCR amplification

From earlier work in our lab, we knew the safest way to perform the reverse transcription (RT) step was to RT a small portion (2 μg) from all of the samples and check the resulting cDNA for it's ability to serve as a template for the PCR of housekeeping genes and to check for any contaminating DNA. cDNA contamination with residual genomic DNA was assessed by primers designed against non-coding sequences from a mouse bac generated in-house (a kind gift from Ruth Young [sts 760 and 761-Sanger notation for sequence tag sites]) (Figure 15). Any contaminated samples were discarded and replaced where possible. Major differences of RNA levels were assessed by use of primers to housekeeping genes. Housekeeping genes are those that are essential to the fully functioning cell, in particular those that are needed for ribosomal activity, and thus presumed to be at a similar level in all tissues. For this exercise, the Ribosomal Protein S29 (Acc.no. L31609) and calcium-binding protein Cab 45b (Acc.no. U45978)(Figure 16) were employed. Primers designed to the ribosomal proteins were used to normalise the samples to account for differences in the amount of total RNA present in each sample and to compensate for any variations in the efficiency of the RT reaction. RNA concentration was assessed by the amount of amplification of these housekeeping genes and only when these assays were satisfactory was a larger amount of RNA committed to the procedure. By introducing this valuable step, it was possible to eliminate samples that were sub-optimal and re-extract better quality RNA. Samples which were replaced due to genomic contamination at this stage included 15.5d forebrain, 17.5d bladder, kidney, and yolk sac. Through carefull examination of the small scale RTPCR, a definitive list of tissues with 50 μg amounts of DNA-free total RNA was compiled and reverse transcribed on a single day with known batches of DNase1 and M-MLV enzymes.

In addition to the 74 mouse foetal tissues and 11 adult mouse tissues, genomic DNA from mouse, rat and human plus glycogen blanks were included in the panel to allow for a direct comparison of known and novel genes.

The next step was to PCR the full range of cDNAs again to recheck for contaminating genomic DNA. Contamination was found in some samples (including 17.5d midbrain and spleen, 15.5d head, rest of body and placenta, 13.5d yolk sac, and 12.5d yolk sac). As we were unable to replace these samples with a genomic free sample, the minimal level of contamination was noted (Figure 15) - these samples were kept in preference to no representation of that tissue sample. This panel was originally designed as a resource to probe with a range of primers designed to within 300bp of the 3' end, a region not thought to contain introns, and so every precaution was taken to eradicate DNA prior to reverse transcription. Levels of cDNA were then adjusted to provide an even amplification of housekeeping genes when PCR'ed at a low cycle number (Figure 16), as assessed by eye from the ethidium bromide stained gels. These calculations are illustrated in the table below as the dilution factor in ul/ml.

1.4.4. Storage of panel

Once an even amplification of housekeeping genes could be demonstrated (see Figure 9) with a minimal degree of genomic contamination (see Figure 15) the panel was judged to be normalised and ready for production. Stocks of cDNAs were stored as 50 □g/ml at –70^{o}C. Working solutions at 2 ng/□l were plated out in 96 well format in a deep well microtitre plate.

Aliquoting was performed with the aide of a Robbins Hydra 96 set to dispense 5 □l into microtitre plates and rapidly frozen on a metal block held on cardice. These plates were then kept at -20^{o}C for up to one month.

Figure 13: Final format for The Mouse Foetal cDNA Panel

tissue order	tube number	ul /ml	gel ID	tissue order	tube number	ul /ml	gel ID
whole conceptuses 8.5d	1	39.5	1	urogenital/gonads 15.5d	46	39.6	49
whole conceptuses 9.5d	2	39.5	2	urogenital/male gonads 17.5d	47	39.6	50
whole foetus 10.5d	3	59.4	3	urogenital/female gonads 17.5d	48	39.6	51
whole foetus 11.5d	4	79.2	4	bladder 15.5d	53	79.2	52
whole foetus 12.5d	90	39.5	5	kidney 15.5d	54	39.6	53
whole foetus 13.5d	91	39.5	6	bladder 17.5d	88	39.6	54
whole foetus 15.5d	92	39.5	7	kidney 17.5d	56	39.6	55
whole foetus 17.5d	93	39.5	8	spleen 17.5d	57	39.6	56
Forebrain 12.5d	5	39.5	9	forelimbs 12.5d	49	39.6	57
Forebrain 13.5d	6	39.5	10	forelimbs 13.5d	50	39.6	58
Forebrain 15.5d	7	39.5	11	forelimbs 15.5d	51	39.6	59
Forebrain 17.5d	8	39.5	12	forelimbs 17.5d	52	39.6	60
midbrain 12.5d	9	39.5	13	rest of body 12.5d	58	39.6	61
midbrain 13.5d	10	39.5	14	rest of body 13.5d	59	39.6	62
midbrain 15.5d	11	39.5	15	rest of body 15.5d	60	59.4	63
midbrain 17.5d	12	39.5	16	rest of body 17.5d	61	59.4	64
Hindbrain 12.5d	13	39.5	17	extra embryonic 10.5d	62	39.6	65
Hindbrain 13.5d	14	39.5	18	extra embryonic 11.5d	63	39.6	66
Hindbrain 15.5d	15	39.5	19	placenta 12.5d	64	39.6	67
Hindbrain 17.5d	16	39.5	20	yolk sac 12.5d	65	39.6	68
spinal cord 12.5d	17	39.5	21	placenta 13.5d	66	39.6	69
spinal cord 13.5d	18	39.5	22	yolk sac 13.5d	67	59.5	70
spinal cord 15.5d	19	39.5	23	placenta 15.5d	68	39.6	71
spinal cord 17.5d	20	59.4	24	yolk sac 15.5d	69	39.6	72
head 12.5d	21	39.5	25	placenta 17.5d	70	39.6	73
head 13.5d	22	39.5	26	yolk sac 17.5d	70b	39.6	74
head 15.5d	23	39.5	27	whole brain - adult	75	39.6	75
head 17.5d	24	79.2	28	spinal cord - adult	76	39.6	76
heart 12.5d	25	79.2	29	skeletal muscle - adult	77	59.4	77
heart 13.5d	26	39.6	30	heart - adult	78	39.6	78
heart 15.5d	27	198	31	liver - adult	79	59.4	79
heart 17.5d	28	39.6	32	kidney - adult	80	79.2	80
lung 12.5d	29	26.4	33	fundus - adult	81	26.4	81
lung 13.5d	30	39.6	34	caecum - adult	82	39.6	82
lung 15.5d	31	59.4	35	testis - adult	83	79.2	83
lung 17.5d	32	39.6	36	ovary - adult	84	52.8	84
liver 12.5d	33	39.6	37	one day old mouse	85	52.8	85
liver 13.5d	34	39.6	38	no RT product	71		86
liver 15.5d	35	59.4	39	Mouse genomic	72		87
liver 17.5d	36	59.4	40	Human genomic	73		88
gut 12.5d	37	39.6	41	Rat genomic	74		89
intestine 13.5d	39	39.6	42	no RT product	71		90
oes/stom 13.5d	40	39.6	43				
intestine 15.5d	41	39.6	44				
oes/stom 15.5d	42	39.6	45				
intestine 17.5d	43b	59.4	46				
urogenital /kidney 12.5d	44	19.8	47				
urogenital/kidney 13.5d	45	59.4	48				

Of the original dissected samples, the following failed to provide sufficient RNA of good enough quality to be included in the panel: 6.5 d, 7.5 d, 15.5 d spleen, and 17.5 d oesophagus/stomach. The 12.5 d gut of embryonic stomach/oesophagus and intestine is a combined sample post RNA extraction. As mentioned earlier, there is included in this list a number of genomically contaminated samples, which we were unable to replace.

1.4.5. Prescreen

A prescreen cDNA panel (plus controls) of whole conceptuses at each of the developmental stages (8) was created to reflect the full range of the panel.

Figure 14: Prescreen

Order	tube number	µl for 1ml	prescreen letter
whole conceptuse 8.5d	1	39.5	A
whole conceptuse 9.5d	2	39.5	B
whole foetus 10.5d	3	59.4	C
whole foetus 11.5d	4	79.2	D
whole foetus 12.5d	90	39.5	E
whole foetus 13.5d	91	39.5	F
whole foetus 15.5d	92	39.5	G
whole foetus 17.5d	93	39.5	H
Glycogen			I
mouse genomic DNA			J

This prescreen panel was employed to check the PCR conditions prior to full profiling. Primers were routinely run at 35 and 45 cycles of PCR with annealing temperatures of 55^{o}C and 60^{o}C. The settings, which allowed the detection of amplicons by ethidium bromide on an agarose gel, at less than saturating conditions, ie before the components of the reaction became limiting and during the time of amplicon accumulation, were generally the most suitable to use for full profiling.

1.5. Panel Validation

Initially a range of specific genes was analysed to demonstrate the integrity and validity of the panel, as compared to data produced with the adult mouse panel.

Sucrase isomaltase (Acc.no. X15546): (Figure 17) an enzyme switched on at birth in response to a diet of lactose in the mother's milk, is expressed exclusively in the adult gut region (Figure 18). As expected, its cDNA was not found in the mouse foetal section of the panel.

Casein, alpha (Acc.no M36780): (Figure 19) was found from the adult panel to be expressed solely in the mammary gland and mid term foetus (Figure 20): interestingly, this gene showed positive in all stages of the head and 17.5 day hindbrain, the 17.5 day midbrain and spleen both showed as positive, but may be due to contamination

Solute Carrier family 17, Na/H exchanger (Acc.no. L33878): (Figure 21) this gene is expressed exclusively in the kidney of the adult panel (Figure 22), and in the foetal panel its cDNA is found in the 17.5 day kidney and spleen samples, this latter finding may be due to the contamination (see figure 15) of the spleen sample.

Calbindin 28K (Acc.No. M23663): (Figure 23) expression in the adult panel is predominantly in the brain, spinal cord, female gonads and kidney (Figure 24) and this pattern is mirrored in the foetal panel, even at the level of 17.5 day uro. + female gonads.

Fatty Acid Binding Protein (Acc.no. M65034): (Figure 25) is expressed exclusively as an intestinal gene in the adult panel (Figure 26), and has strong expression in a similar region in the foetal panel.

Other genes, which had shown very specific patterns in adult mouse tissues, were less easy to interpret when run on the mouse foetal panel. These include ***Adora 1*** (Figure 27) and ***Hox b5*** (Figure 29). Figures for mouse adult panels of these genes are

included (28 and 30). The *Adora1* adult profile illustrates expression to be widely found in the brain and gonads; with the foetal panel, expression is found in parts of the brain, developing gonads and interestingly also the gut regions. This may be a result of contamination during the processing of this tissue or a true result for this gene. *Hoxb5* shows clear expression in the region of the adult cerebellum, spinal cord and gut regions; in the foetal panel, expression is found in the mid/hind brain, spinal cord, gut regions, and developing gonads. Expression found in the midbrain is clearly a result of contamination as the gene *Hoxb5* is well documented as a gene expressed in the upper rhombomers only during developement, this would include the hindbrain only. Expression also found in the 17.5 day spleen (see figure 15) is due to contamination.

In order to thoroughly examine the band intensities of a selection of the gels, images were photographed directly from the UV light box, relayed to a computer and the specific band intensities recorded. Ethidium bromide, the commonly used molecular biology stain, at a concentration of 2ug/ml forms a complex with DNA. The fluorescence emission for the DNA-ethidium bromide complex peaks at 300nm and is between 10-50 times that of the uncomplexed dye. With a UV transilluminator light source of 302nm, the DNA-ethidium bromide complex fluoresces in proportion to the amount of complex present [1]. Through the use of densitometric measurements, the digital image of the fluorescence was measured in a similar way to commercially available software http://www.uvp.com/. The lane containing the band from mouse genomic DNA was used as a reference and given a score of 100, to which all other bands measurements were related to. It should be noted that the fluorescent intensities of the ethidium bromide-DNA complexes can be reduced by as much as 40% over a 10 minute period and so it was important that image capturing was performed in a timely manner. In this way the profiles for housekeeping genes from

Figure 16 are displayed graphically as Figure 31 to illustrate the level of evenness obtained through normalisation and to show how these two genes vary in their expression patterns (Figure 35 is included in this chapter to allow identification of tissues). The *Ribosomal Protein 29* is the profile labelled 310 and *Stromal Cell derived factor 4 (Sdf4)* (previously known as *Cab45 [2])* is the profile labelled 428. Ribosomal proteins are regarded as genes essential for the cell's maintenance as they are required for translation. *Stromal cell derived factor 4* is a calcium ion protein, found to be expressed in all adult mouse tissues [3]. Use of this approach highlighted the inaccuracies undetected by eye from the ethidium bromide stained gels, in particular it demonstrated the lack of concordance between the two presumptive housekeeping genes and the inter-tissue variations. The agarose gels displayed have expression levels marked strong through to trace, presenting a broad range of scoring.

1.6. Concluding Remarks

It was through validating the panel against these genes (Figures 17 – 29), which had shown reproducibly specific expressions in the adult panel, that a representation of adult mouse tissues was subsequently included in the final panel format (Figures 15 and 16). Human and rat genomic samples were also included to allow inter species comparison. The panel was considered sufficiently robust to continue with the test by applying primers to the family of *Sox* genes.

Legend for Figures 15 through to 30

Ethidium bromide stained 2.5% agarose gels, run at 125 volts, 30 milliamps, 2 hours and photographed with a Polaroid camera. Photographs were scanned on an Epson scanner, opened in Adobe Photoshop, where defined regions were then cut and pasted into MacDraw and scored through viewing the original photographs.

Gene/Est: Gene or EST name
Symbol: Acronym for gene/est
EMBL AcNo: Accession number from EMBL database
Primers: Sanger Centre identifying number

PCR: Polymerase Chain Reaction
Date: date of lab. work
$MgCl_2$: final concentration
Temp: annealing temperature
Cycles: number of cycles

Amplification signal: four catagories
Strong – dark blue – major signal from both duplicates.
Moderate-weak – blue – broadest range of signal.
Trace – green – faint signal.
None – yellow – no signal from either duplicate.
Discrepancy – red – signal from one only, of the duplicates.

Figure 15: Contamination Profile

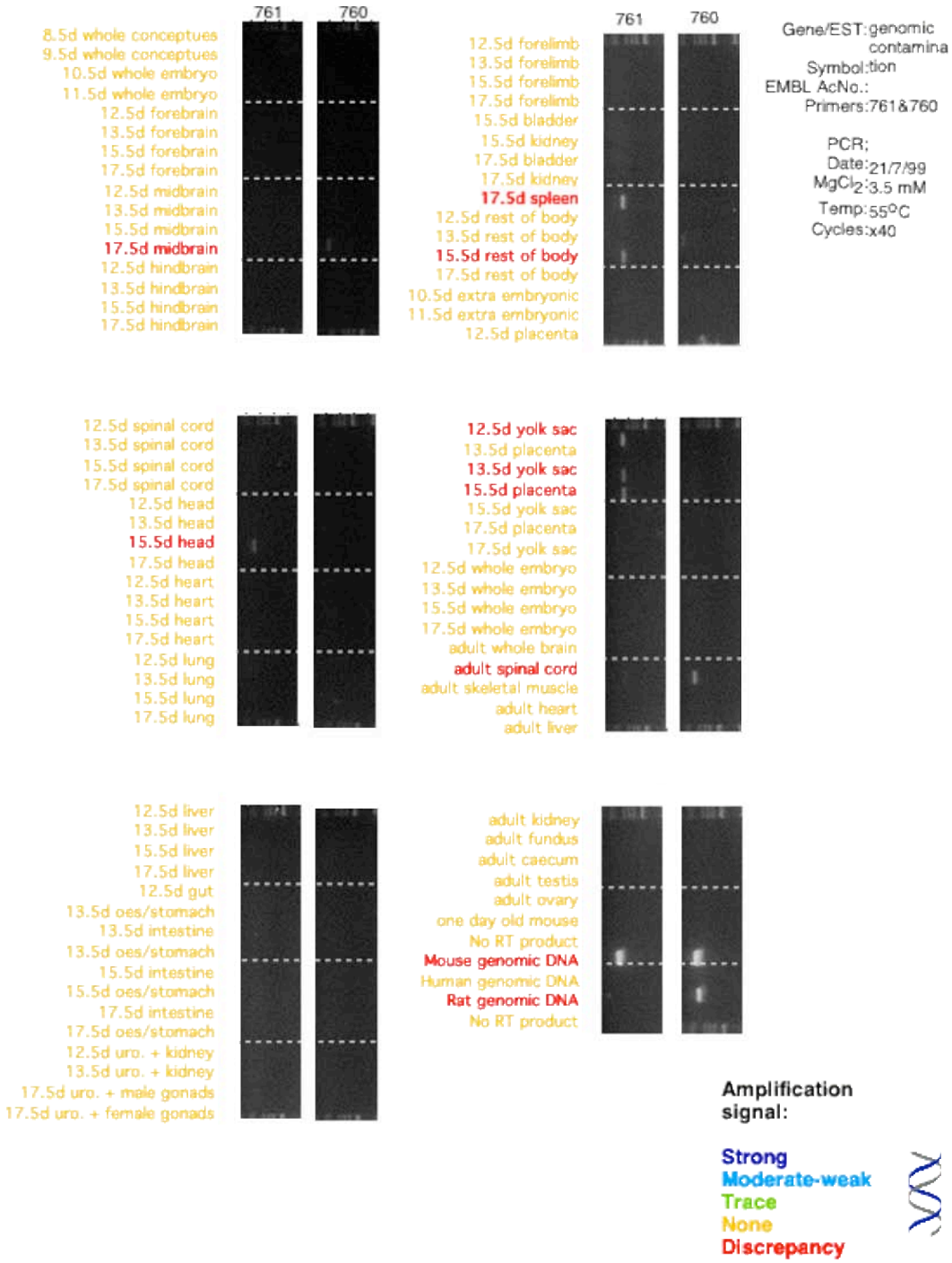


Figure 9: Normalisation Profile

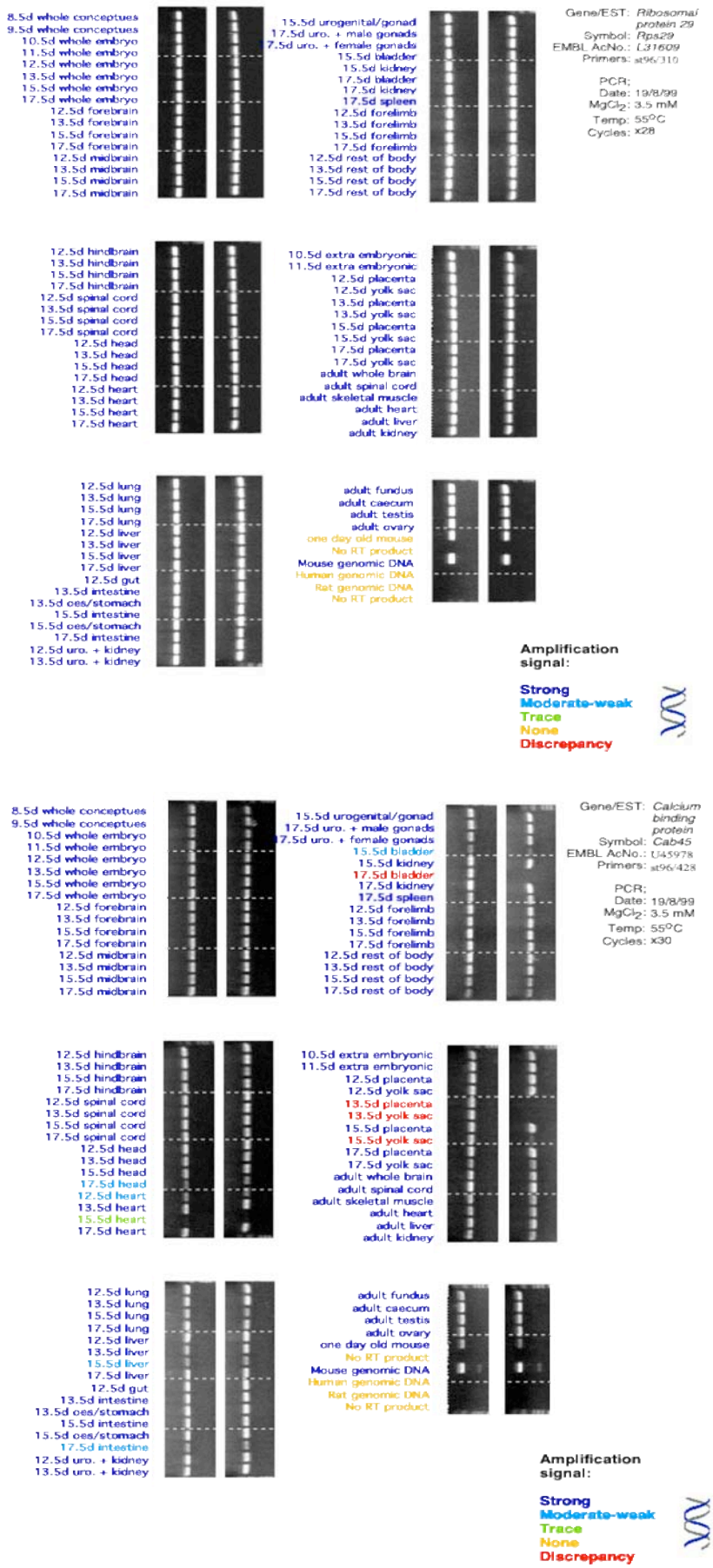


Figure 17: Expression Profile of Sucrase Isomaltase in Mouse Foetal Panel

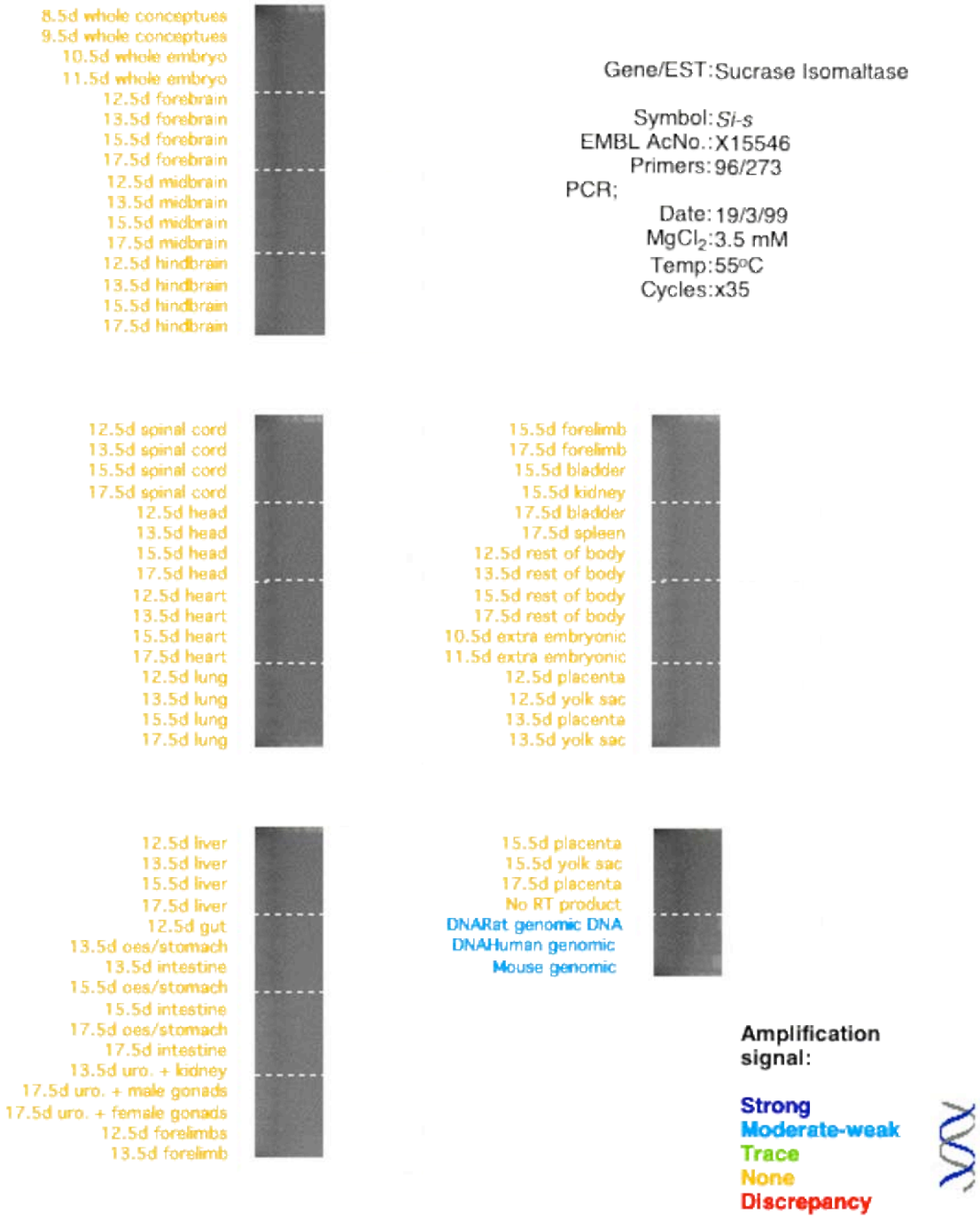


Figure 18: Expression Profile of Sucrase Isomaltase in Mouse Adult Panel

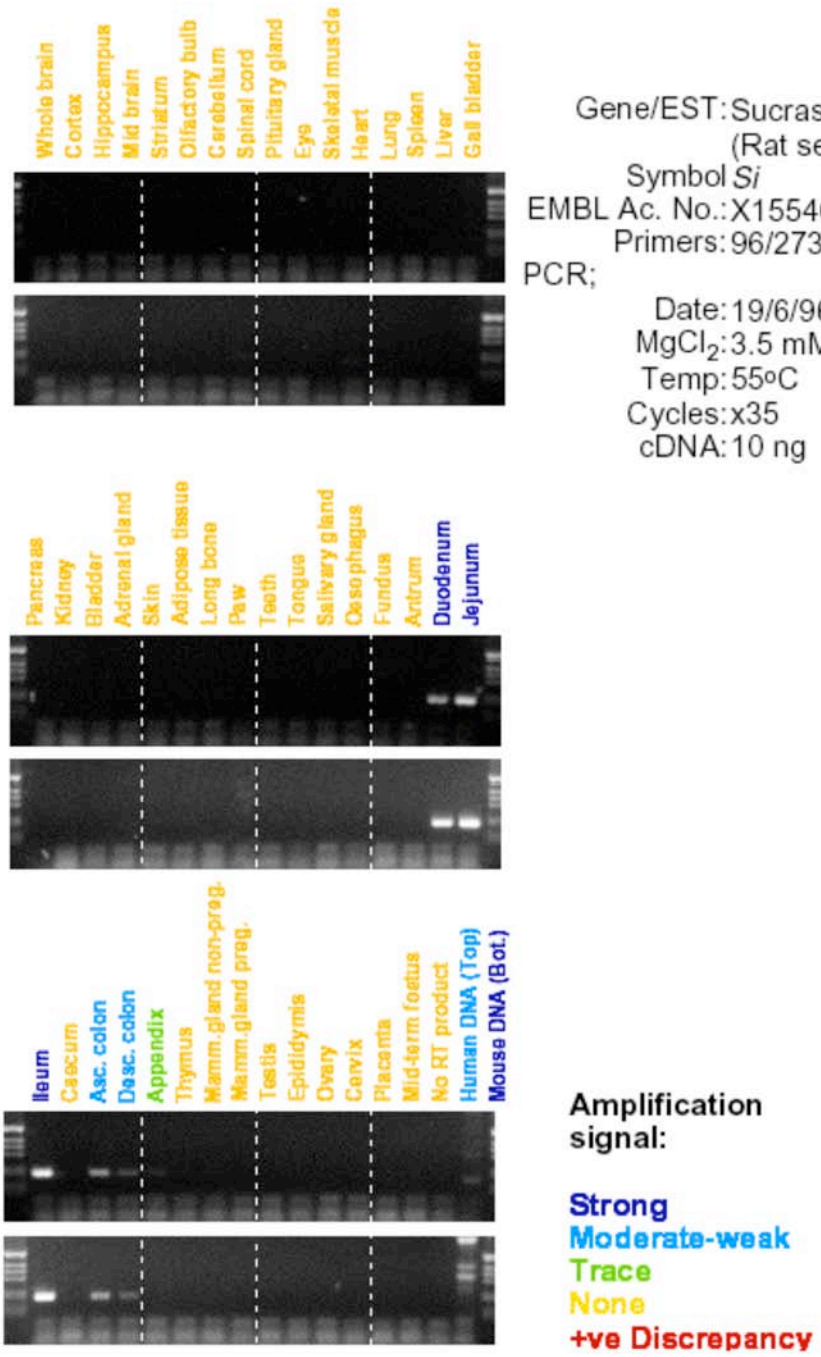


Figure 19: Expression Profile of Casein, alpha in Mouse Foetal Panel

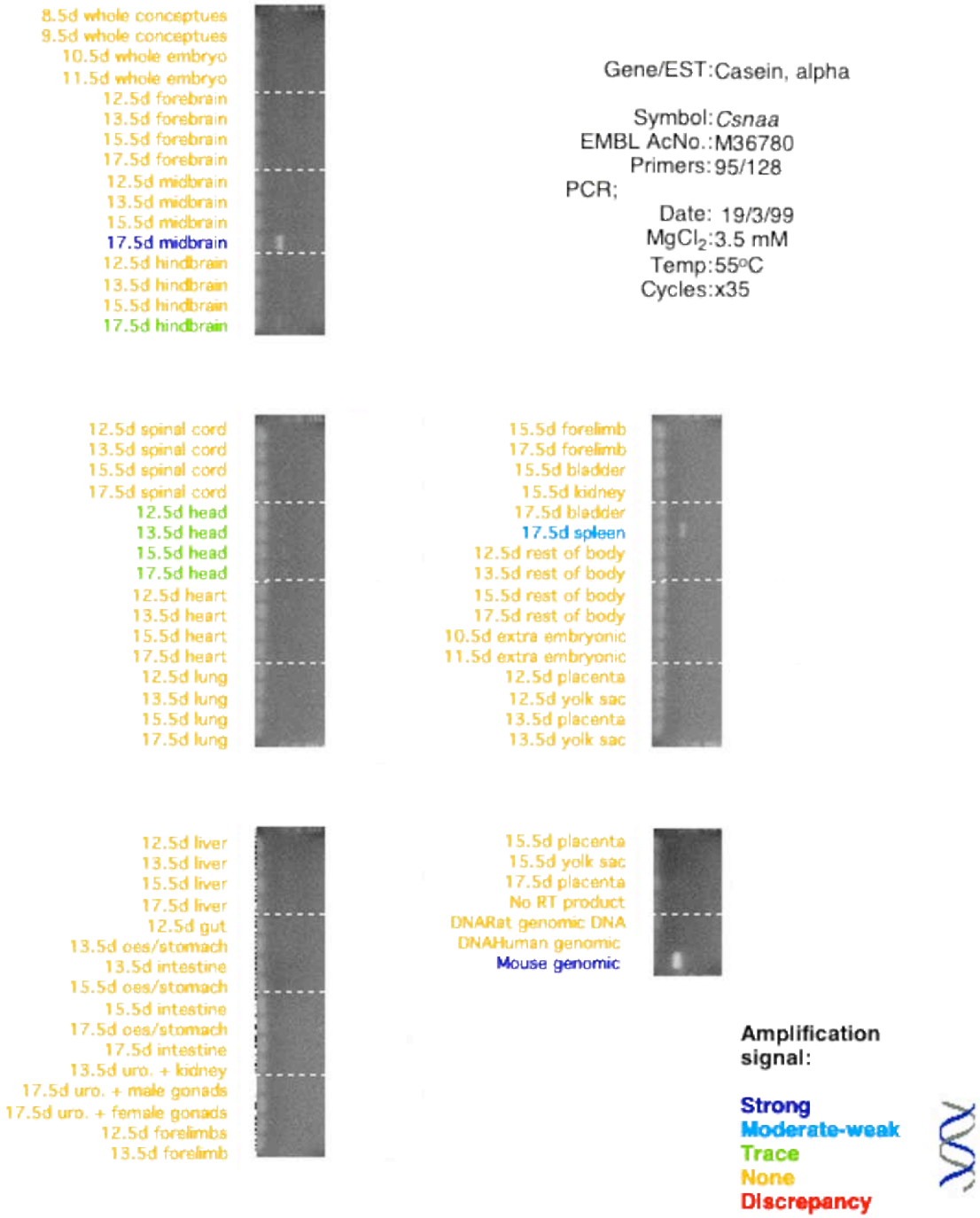
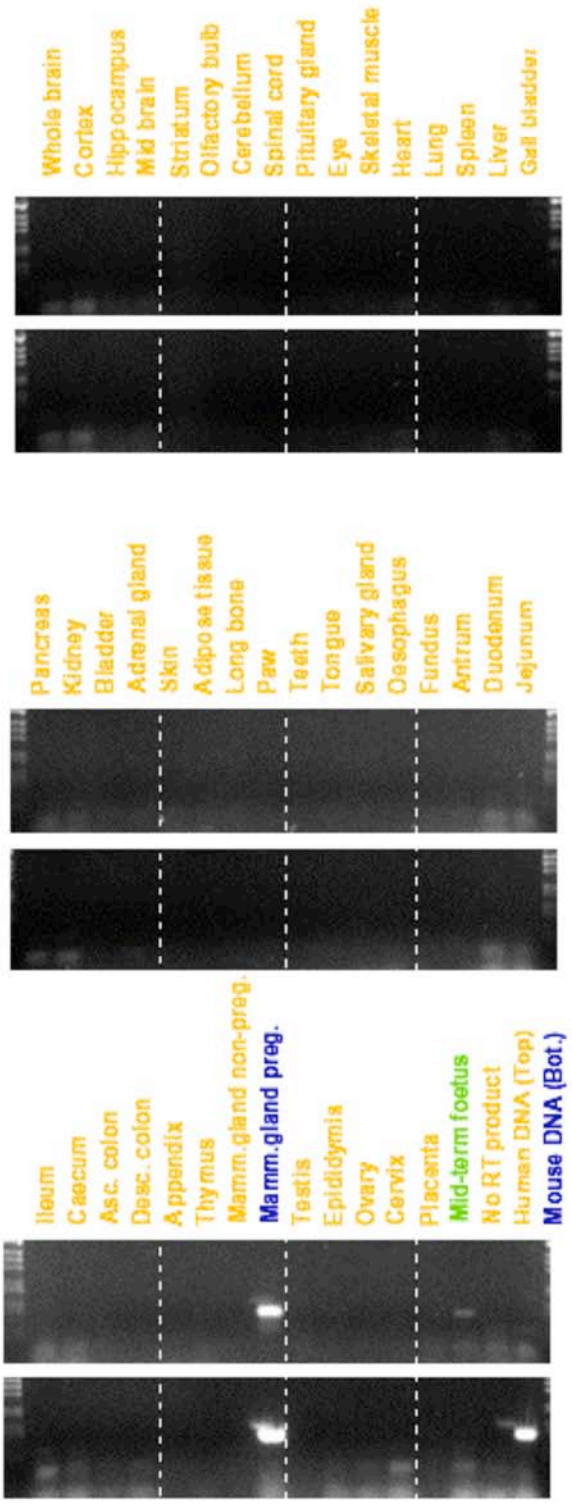


Figure 20: Expression Profile of Casein, alpha in Mouse Adult Panel



Gene/EST: Casein alpha
Symbol: *Csna*
EMBL Ac. No.: M36780
Primers: 96/128
PCR;
Date: 3/4/96
$MgCl_2$: 3.5 mM
Temp: 55°C
Cycles: x 35
cDNA: 10 ng

Amplification signal:

Strong
Moderate-weak
Trace
None
+ve Discrepancy

Figure 21: Expression Profile of Solute Carrier 17a2 in Mouse Foetal Panel

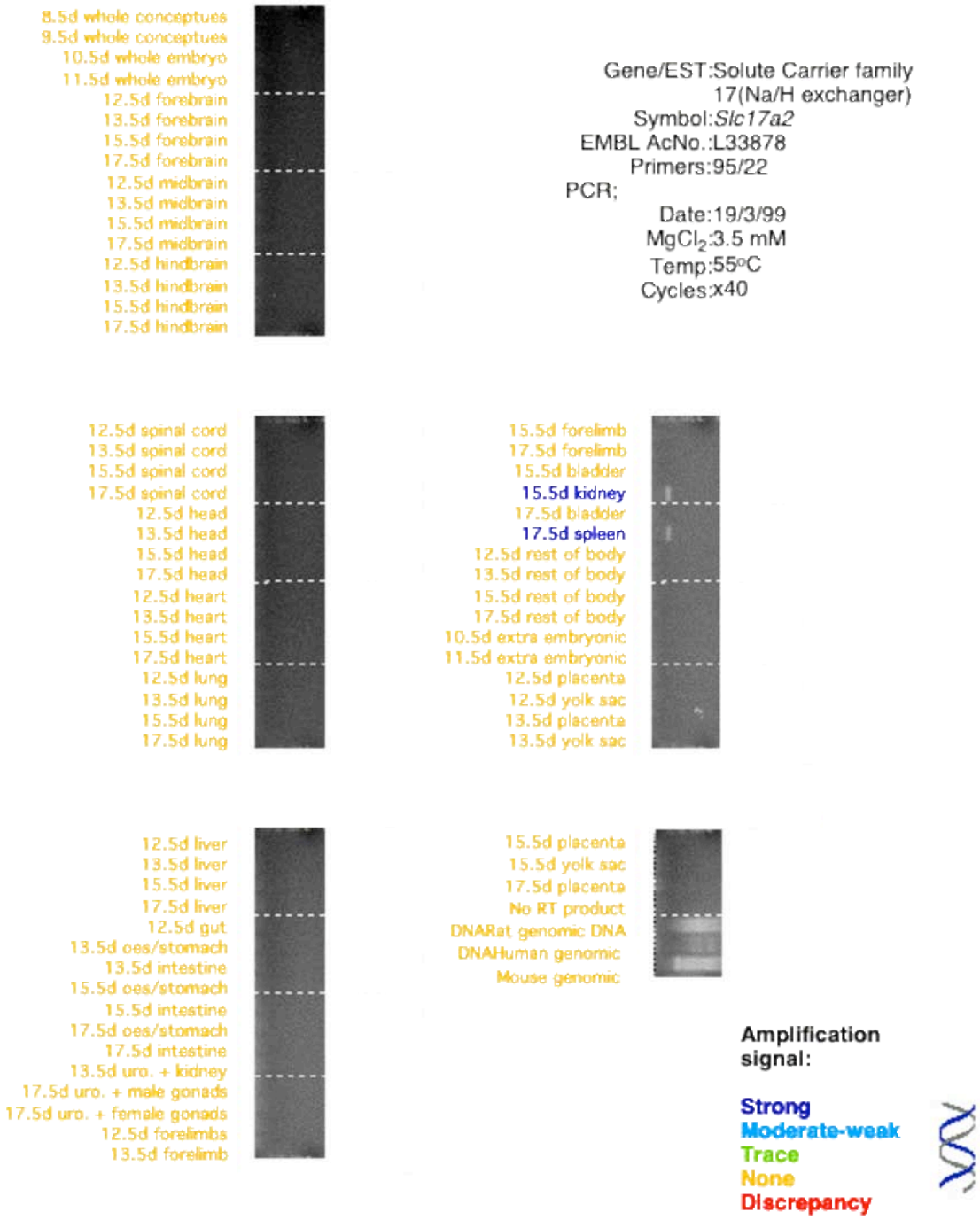
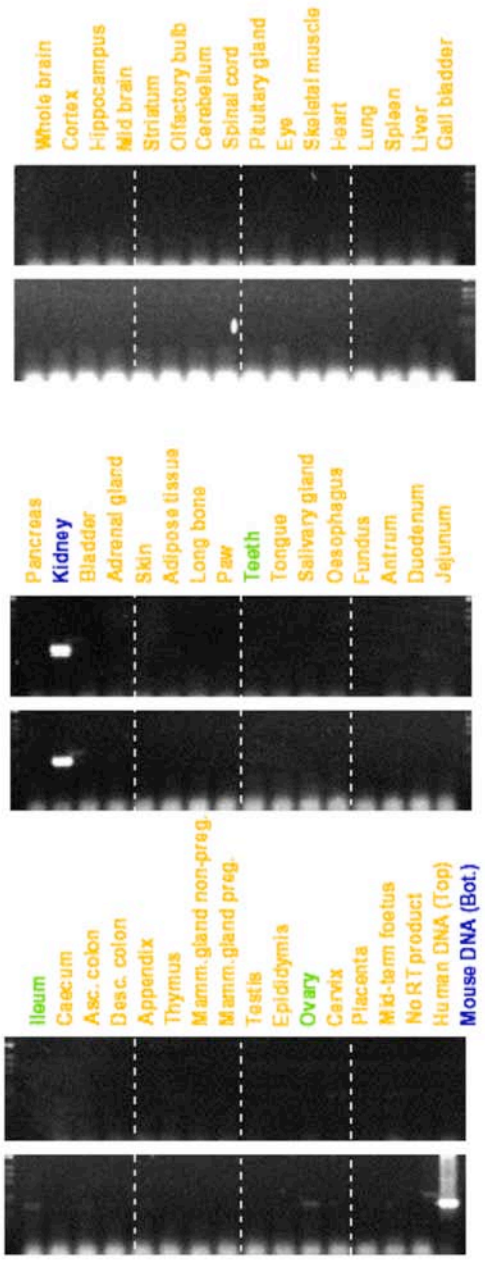


Figure 22: Expression Profile of Solute Carrier 17a2 in Mouse Adult Panel



Gene/EST: Solute carrier family 17 (sodium/hydrogen exchanger) member 2
Symbol: *Slc17a2*
EMBL Ac. No.: L33878
Primers: 95/22
PCR;
Date: 8/2/96
$MgCl_2$: 3.5 mM
Temp: 55°C
Cycles: x40ex

Amplification signal:

Strong
Moderate-weak
Trace
None
+ve Discrepancy

Figure 23: Expression Profile of Calbindin 28K in Mouse Foetal Panel

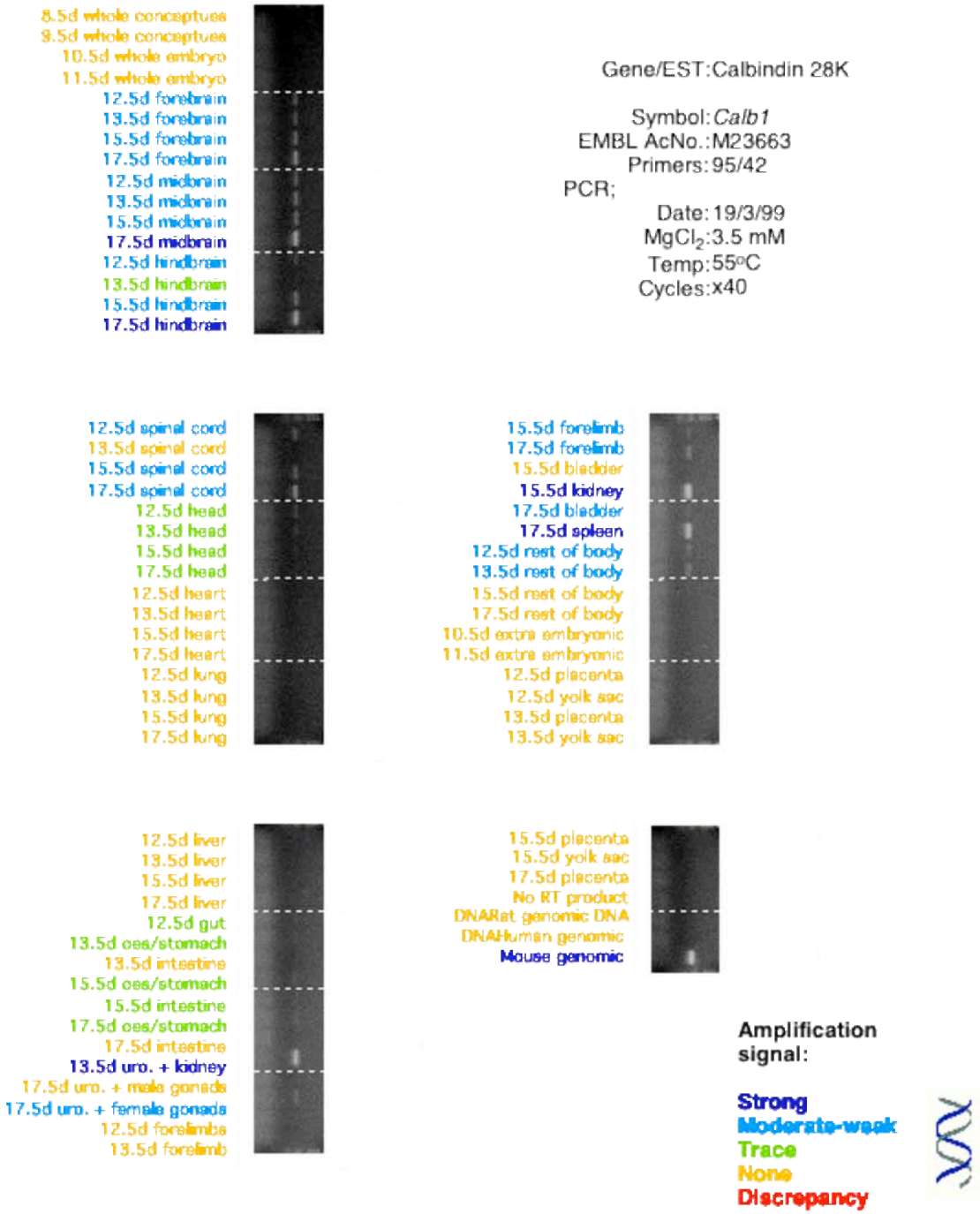
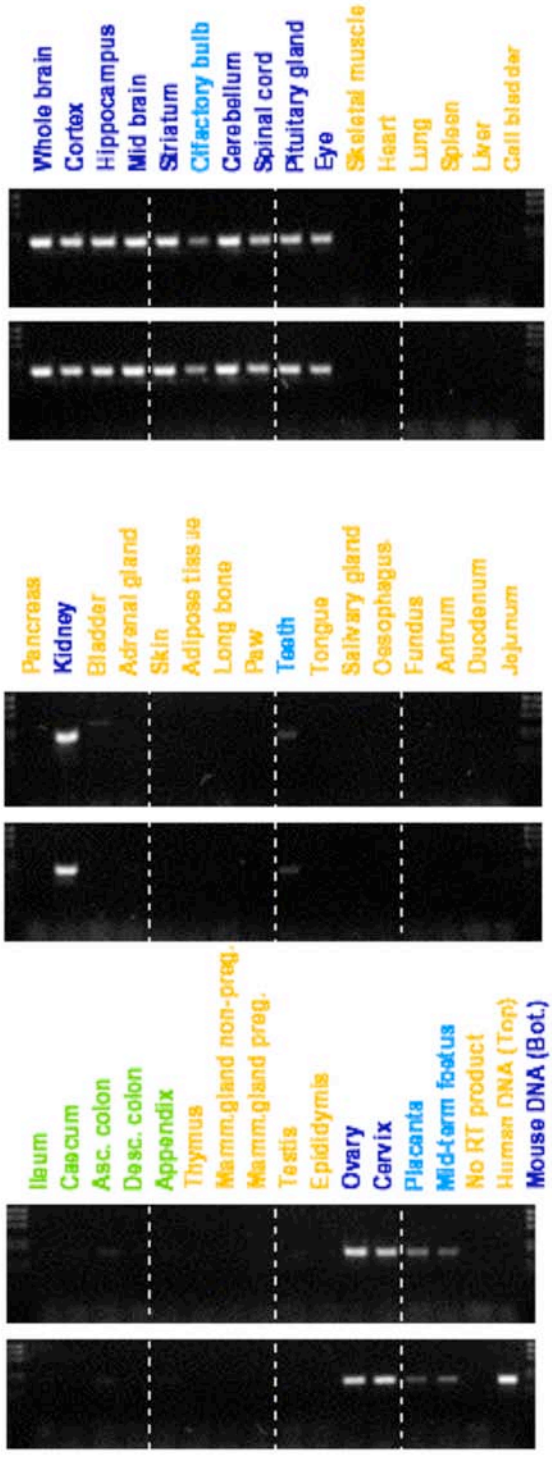


Figure 24: Expression Profile of Calbindin 28K in Mouse Adult Panel



Gene/EST: Calbindin-28K
Symbol: *Calb1*
EMBL Ac. No.: M21531
Primers: 95/35
PCR;
Date: 31/1/96
$MgCl_2$: 3.5 mM
Temp: 55°C
Cycles: x50ex

Amplification signal:

Strong
Moderate-weak
Trace
None
+ve Discrepancy

Figure 25: Expression Profile of Fatty Acid Binding Protein in Mouse Foetal Panel

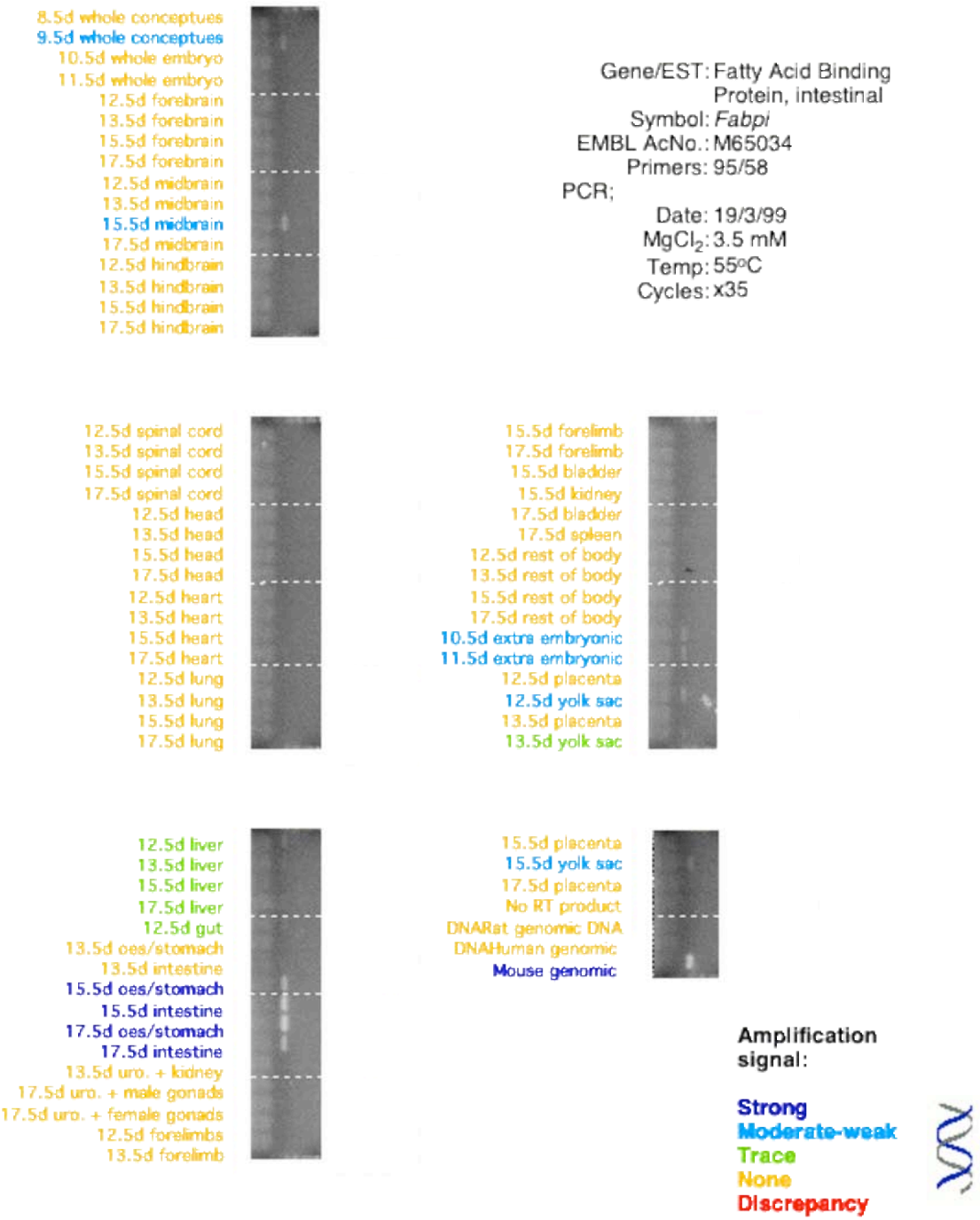


Figure 25: Expression Profile of Fatty Acid Binding Protein in Mouse Adult Panel

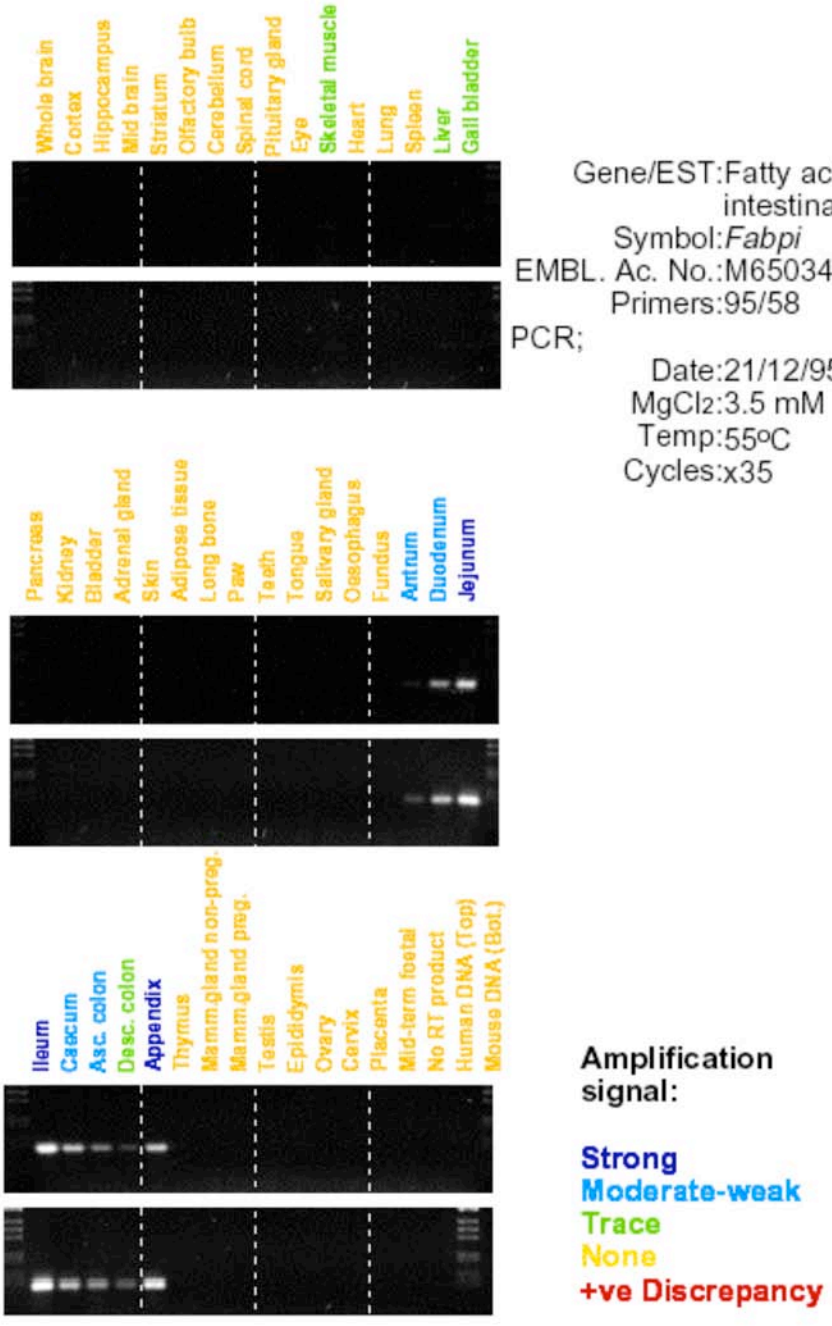


Figure 27: Expression Profile of Adenosine A1 in Mouse Foetal Panel

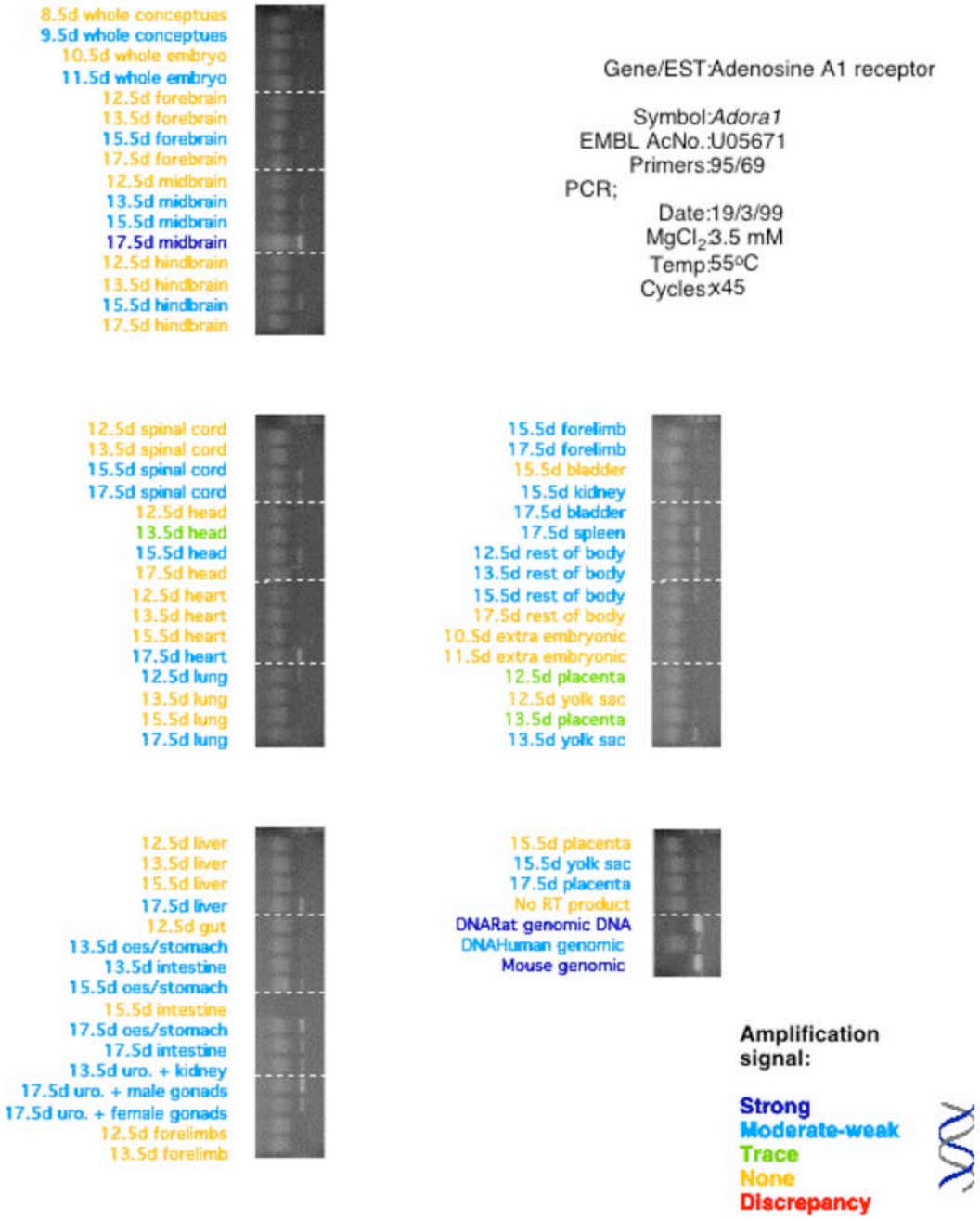
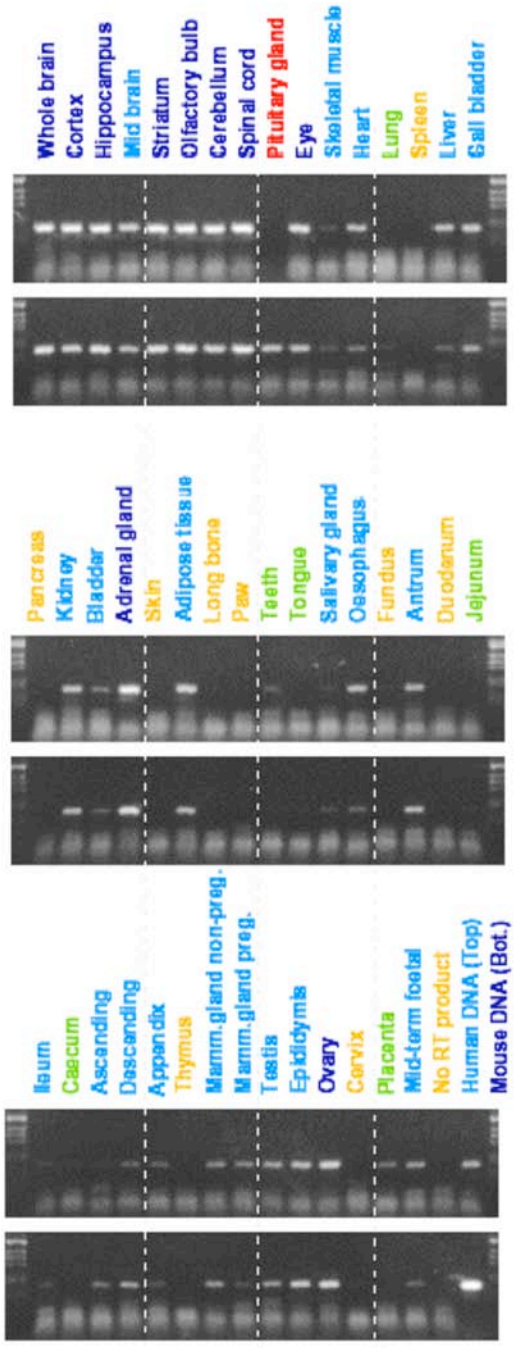


Figure 28: Expression Profile of Adenosine A1 in Mouse Adult Panel



Gene/EST: Adenosine A1 receptor
Symbol: *Adora1*
EMBL Ac. No.: U05671
Primers: 95/69
PCR;
Date: 14/3/96
$MgCl_2$: 3.5 mM
Temp: 55°C
Cycles: x45ex

Amplification signal:

Strong
Moderate-weak
Trace
None
+ve Discrepancy

Figure 29: Expression Profile of Homeo box B5 in Mouse Foetal Panel

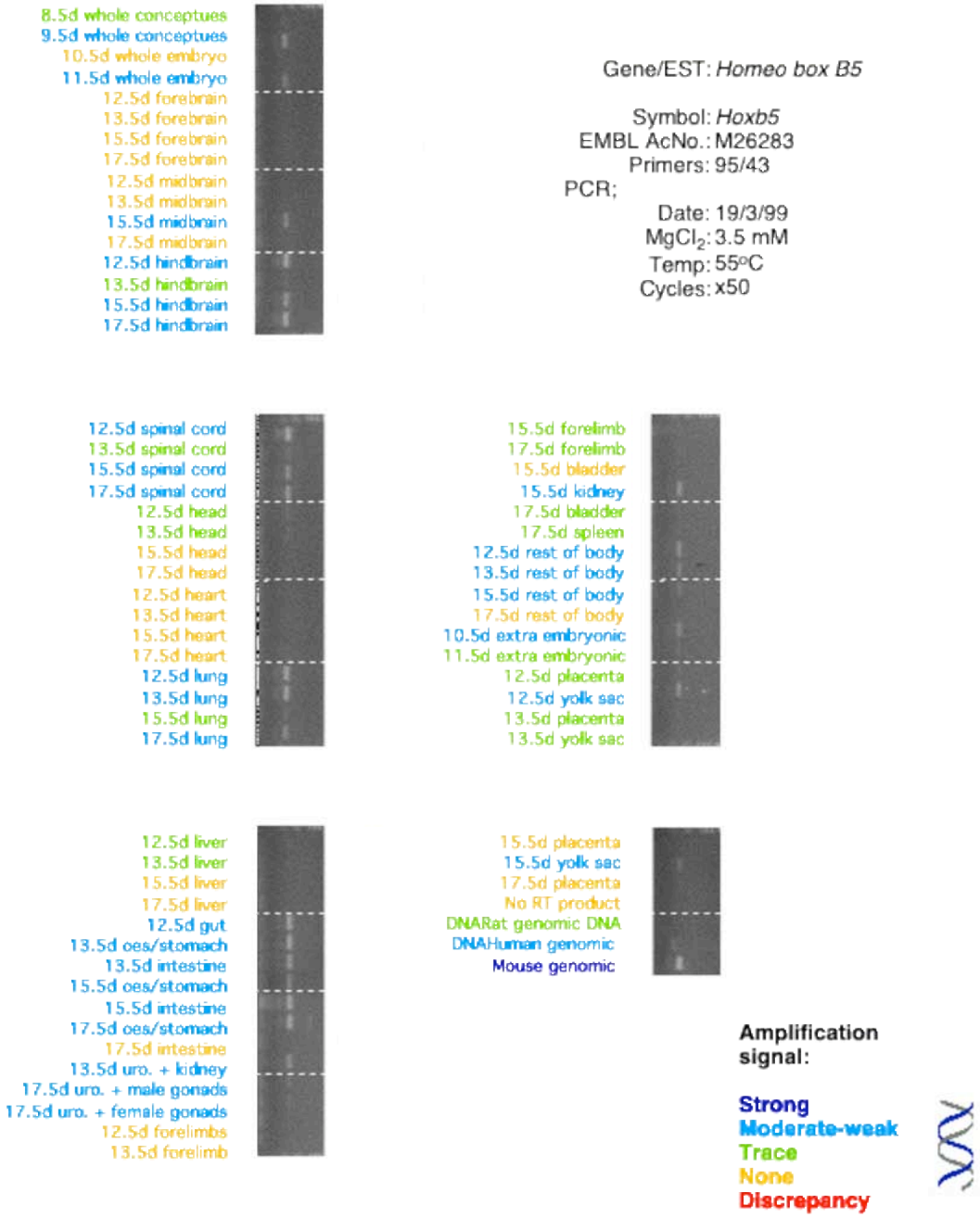
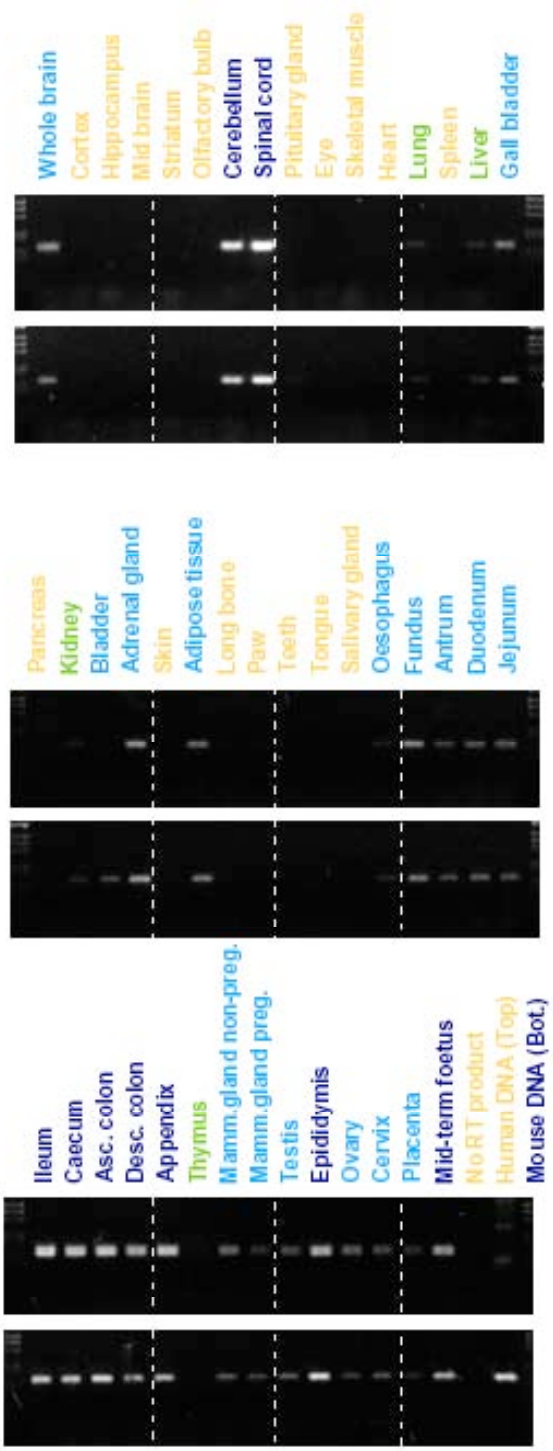


Figure 30: Expression Profile of Homeo box B5 in Mouse Adult Panel



Gene/EST:Homeo box B5
Symbol:*Hoxb5*
EMBL Ac. No.:M26283
Primers:95/43
PCR;
Date:31/1/96
$MgCl_2$:3.5 mM
Temp:55°C
Cycles:x50ex

Amplification signal:

Strong
Moderate-weak
Trace
None
+ve Discrepancy

Figure 31: Graphical representation of gel images in Figure 16

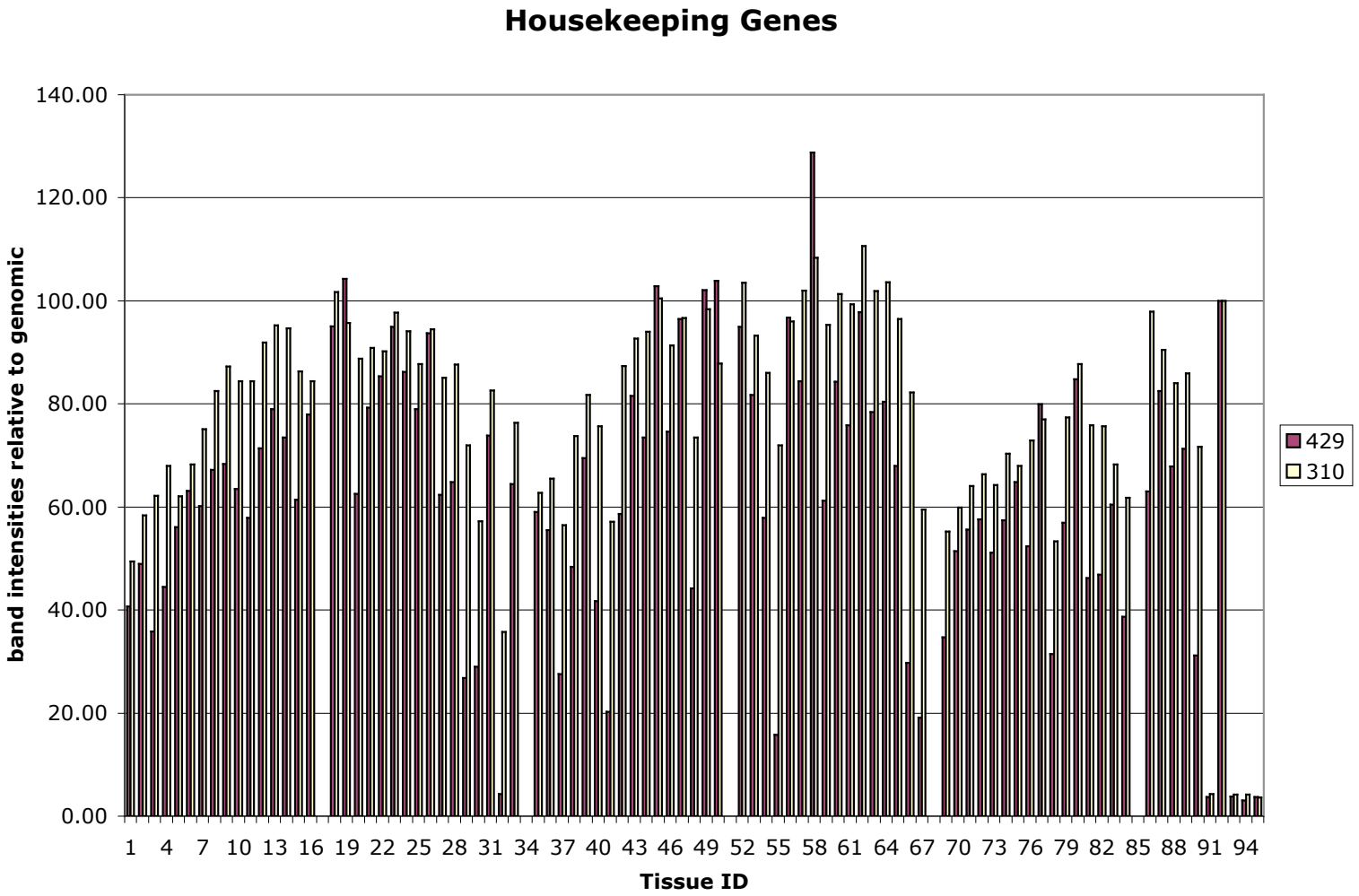


Figure 35: Tissue identification of graph position

ID	tissue
1	whole conceptuses 8.5d
2	whole conceptuses 9.5d
3	whole embryo 10.5d
4	whole embryo 11.5d
5	whole embryo 12.5d
6	whole embryo 13.5d
7	whole embryo 15.5d
8	whole embryo 17.5d
9	forebrain 12.5d
10	forebrain 13.5d
11	forebrain 15.5d
12	forebrain 17.5d
13	midbrain 12.5d
14	midbrain 13.5d
15	midbrain 15.5d
16	midbrain 17.5d
18	hindbrain 12.5d
19	hindbrain 13.5d
20	hindbrain 15.5d
21	hindbrain 17.5d
22	spinal cord 12.5d
23	spinal cord 13.5d
24	spinal cord 15.5d
25	spinal cord 17.5d
26	head 12.5d
27	head 13.5d
28	head 15.5d
29	head 17.5d
30	heart 12.5d
31	heart 13.5d
32	heart 15.5d
33	heart 17.5d
35	lung 12.5d
36	lung 13.5d
37	lung 15.5d
38	lung 17.5d
39	liver 12.5d
40	liver 13.5d
41	liver 15.5d
42	liver 17.5d
43	gut 12.5d
44	intestine 13.5d
45	oes/stom 13.5d
46	intestine 15.5d
47	oes/stom 15.5d
48	intestine 17.5d
49	urogenital/kidney 12.5d
50	urogenital /kidney 13.5d

ID	tissue
52	urogenital/gonads 15.5d
53	urogenital male gonads 17.5d
54	urogenital female 17.5d
55	bladder 15.5d
56	kidney 15.5d
57	bladder 17.5d
58	kidney 17.5d
59	spleen 17.5d
60	forelimbs 12.5d
61	forelimbs 13.5d
62	forelimbs 15.5d
63	forelimbs 17.5d
64	rest of body 12.5d
65	rest of body 13.5d
66	rest of body 15.5d
67	rest of body 17.5d
69	extra embryonic 10.5d
70	extra embryonic 11.5d
71	placenta 12.5d
72	yolk sac 12.5d
73	placenta 13.5d
74	yolk sac 13.5d
75	placenta 15.5d
76	yolk sac 15.5d
77	placenta 17.5d
78	yolk sac 17.5d
79	whole brain - adult
80	spinal cord - adult
81	skeletal muscle - adult
82	heart - adult
83	liver - adult
84	kidney - adult
86	fundus - adult
87	caecum - adult
88	testis - adult
89	ovary - adult
90	one day old mouse
91	glycogen
92	mouse genomic
93	human genomic
94	rat genomic
90	glycogen

References

1. Prunell, A., *A photographic method to quantitate DNA in gel electrophoresis*. Methods in Enzymology, 1980. **65**(1): p. 353-358.
2. Ringwald M, et al., *The mouse gene expression database*. Nucleic Acids Research, 2001. **29**: p. 98 - 101.
3. Freeman, T.C., et al., *Expression Mapping of Mouse Genes*. MGI Direct Data Submission, 1998.

CHAPTER 2 Profiling of the Family of Sox Genes by PCR

Introduction

The mouse foetal cDNA panel created in Chapter I was then tested by PCR with primers for members of the *Sox* gene family. A selection of these *Sox* genes, which showed differing expression patterns, was then taken for analysis by *in-situ* hybridization, as described in Chapter 3.

2.1 Special Equipment and Suppliers

Computer	Sun workstation
Computer	Apple Mac.
Software	Mac Draw, Adobe photoshop, Microsft Office, Blixem, AceDB, Internet Explora

2.2 Materials and Solutions

As described in Chapter 1

2.3 Establishment of the *Sox*-related sequences for study

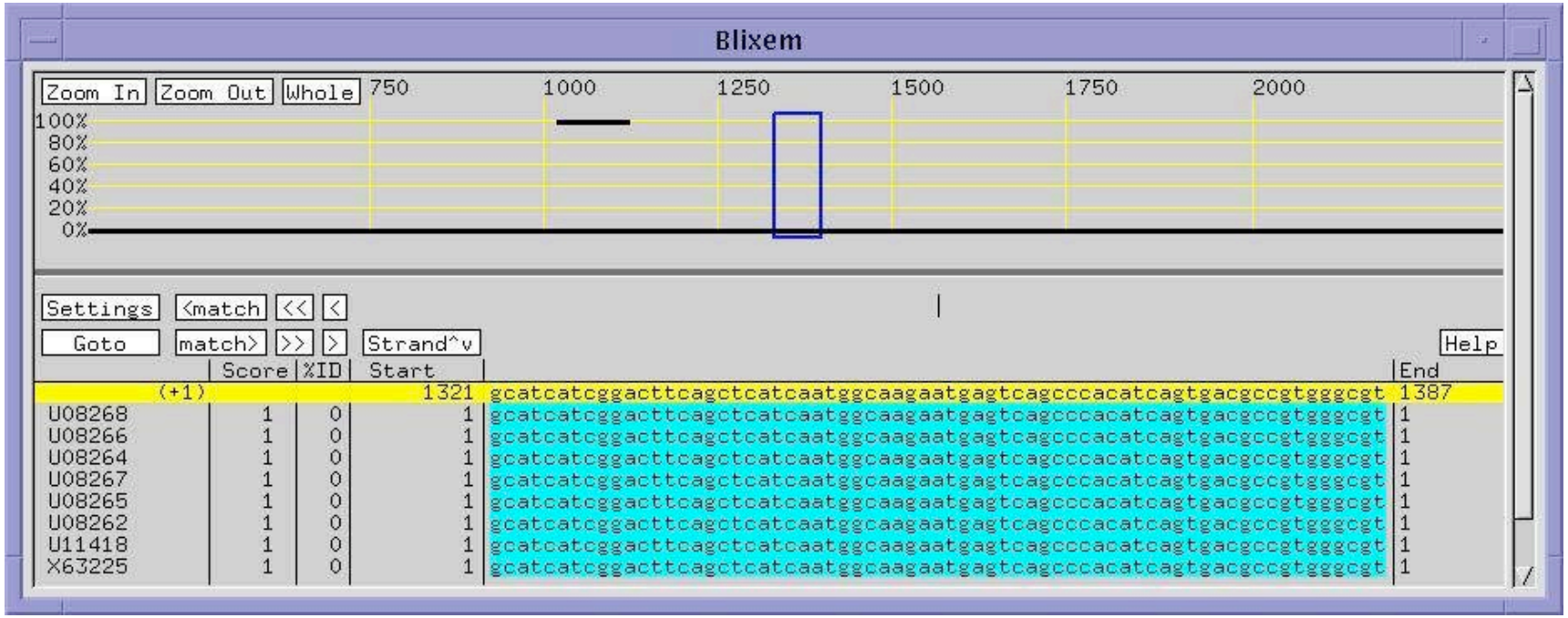
All known *Sox*-related sequences were obtained from the embl website and interrogated for primer design. This was done using the AceDB software designed by Durbin and Thierry-Mieg (1991; http://www.sanger.ac.uk/Software/Acedb/) and adapted as a repository for *Sox*-related sequences. Using the query form from http://www.sanger.ac.uk/SRS and the fields 'organism' and 'all text', for *Mus musculus* and Sox, respectively, a list of sequencing information was imported from

the embl and embl-new databases into this "Soxace database", with the kind assistance of Adam Butler.

It was necessary to determine how many of these genes were represented in this dataset. Therefore, the 274 collected sequences were clustered using the Clustal W and Phrap packages to ensure a 98% homology over a 100 base-pair stretch of sequence, resulting in 16 distinct cluster groups. Tight clustering stringency ensured that each of the *Sox* genes separated into different groups. The *Sry* group held 28 sequences, *Sox 2, 4, 5, 6, 7, 9, 11, 13* and *18* all contained a number of sequences. *Sox 1, 3, 8, 14, 16* and *19*, together with 63 other *Sox*-related sequences failed to group during this exercise. This study highlighted the current interest in the *Sox* genes as illustrated by the cluster sizes. Interestingly, at the time of data collection (July 1999), the *Sox 4* related gene showed the greatest number with 43 and 48 sequences in two separate groups. There are currently 687 *Sox*-related sequences held in the embl database (November 2003).

Blixem software ("BLast matches In an X-window Embedded Multiple alignment") was used to examine the clustered sequences, see Figure 32. This software was developed at the Sanger Institute by Dr. R. Durbin and provides a graphics window, to view sequence alignments of up to 1.2 megabases [1]. Those sequences with greatest homology to all others in their cluster, incorporating the longest 3' region, were chosen for primer design (section 1.3.10).

Figure 32: Blixem Software



Blixem software used to view sequence alignment and identify similarities at the nucleotide level.

Of the primer pairs designed, seven (for *Sox 3, 13, 16, 17, 18, LZ* and *Sry*) failed at the prescreen stage (1.4.5.) and were redesigned. Despite a number of trials at prescreen, we failed to obtain successful primer pairs for *Sry, Sox 3* and *Sox18*. A list of all the primers used and the expected amplicon product sizes are in Figure 33 with further details in Figure 34 [NOTE: the final primer sets that failed the prescreen for *Sry, Sox 3* and *Sox18* are also included in brackets]. Failed PCRs were generally attributed to primer design, after one redesign of primers were usually designated as fails and no longer pursued.

Figure 33: Expected amplicon sizes for PCR

Sox 1	224
Sox2	184
[Sox3	211]
Sox4	210
Sox5	239
Sox6	238
Sox7	113
Sox8	106
Sox9	104
Sox11	151
Sox12	151
Sox13	333
Sox14	100
Sox15	198
Sox16	151
Sox17	175
[Sox18	278]

Sox19	164
Sox21(Sox10)	203
SoxLZ	327
Slc17a2	175
Hox5b	164
Fsbpi	126
Adora1	179
Csna	128
Fabph1	157
Si-S	203
Calb1	187
Rsp29	134
Cab45	170
[Sry]	
bac 573K1b	265
bac 573K1a	261

Figure 34: Primer Sequences

Sanger ID	Forward Primer	Reverse Primer	Symbol	Accession number
St95_22	ATGTGTGTGTGTGTGCACATG	TACACCCGCACTAATGGTCA	Slc17a2	L33878
st95_35	GCTTCTATCTGGCGGAAGG	TGTCATCTGGCTACCTTCCC	Calb1	M21531
st95_42	GCTTCTATCTGGCGGAAGG	TGTCATCTGGCTACCTTCCC	Calb1	M23663
st95_43	CTCTTTTCTCCCACCTCATCC	CTACCAGGCAGCAGGAGTTC	Hox5b	M26283
st95_58	AAGTCAACTTCTCAGAGCCTGG	GCTTTGACAAGGCTGGAGAC	Fsbpi	M65034
st95_69	CCCAGAAGTACTACGGGAAGG	CGAGTTGCCGTGTGTGAG	Adora1	U05671
st96_128	GCCAATGATTCATCTTGAGTTG	CCTTGATTCTCTCCGCTCAG	Csna	M36780
st96_153	CTCATGGTTTTCCCCTCTGA	GGTTCTGCTTTATTGACCTTGG	Fabph1	U02883
st96_273	GGGGAAGTGGAACACACGG	AGCAGGAGTTGGCTGGAATG	Si-s	X15546
st96_310	CTGATCCGCAAATACGGG	GCATGATCGGTTCCACTTG	Rps29	L31609
st96_428	GAAGAGTTCTGAGCATGCCC	TTCTTGGGGCCTATGGAAG	Cab45	U45978
st97_595	AAAAAAAAATGCCCATGCAG	TACGGAAAATAAAAGGGGGG	Sox2	U31967
st97_596	TATGGTATGAAGATGGACGGC	CATGCGGGCTCTTTAAGAAC	Sox6	U32614
st97_613	TTGAAGAAGCCCTGTCCG	TGGCACTGTTTAACCCATAGC	Sox5	X65657
st97_614	ACGTCTGGCAGTGCAGAAC	CTGCCTCATCCACATAGGGT	Sox7	X65660
st97_618	TTTCAGCTCCTCATCGGC	CTCCTCTCCTGCCTCTTGG	Sox4	X70298
st97_625	AATCCCCTCTCAGACGGTG	TTGATGCATTTTGGGGGTAT	Sox1	X94126
st97_627	GTACACAACTACCCAAGGGAGC	GTGGAAGAGTCTGGGGATAGG	Sox15	X98369
st97_629	TAACGCAGAGCTCAGCAAGA	CTTGTGCTGCACACGGAG	Sox8	Z18957
st97_630	AGGAAGCTGGCAGACCAGTA	ACGAAGGGTCTCTTCTCGCT	Sox9	Z18958
st97_632	CGAGCGCAGGAAGATCAT	ATCAGCCATGTGCTTGAGG	Sox11	Z18960
st97_635	GATGGCCCAGGAAAATCC	GATGTAAGGCCGCTTCTCTG	Sox14	Z18963
st97_760	ATGGCTTGATTGGTACCAGTG	GACAAGTGGAAAAACAGGAAGC	bac 573K1a	
st97_761	TGCAGGCAGAGATGCTACTG	CGCTCAGAGAGAAAAAAATTGG	bac 573K1b	
st97_787	TCAAGGGAAAAGAAGAGGGC	TTCAGCGTATCCACCACATCG	SoxLZ	D61689
st97_785	CACGGAGCCATTTTTAGCTC	TAGGGTCAACAGCGGTGA	Sox13	AJ000740
st97_786	TGTCTGCCACTTGAACAGTTG	TGAGAAAACACGCATGACAA	Sox17	D49473
st97_788	AGGCTGGACACTAAACCCCT	AAGTTAATCGGGGCTGGAGT	Sox21	D87031
[st97_789	GGATGACTGACTGGCCATCT	GGCTGCTATTTTCTTCCAACC	Sry	E11536]
st97_790	CATGGTGTGGAGCTCTGCTC	ACGAAGACGCTTAGCCTCCT	Sox16	L29084
[st97_791	CAGGCTAGACACTGTCCTTGC	GAACATAAATGGCAGAAAAGCC	Sox18	L35032]
[st97_792	ATGGGCTCCGTGGTGAAGT	TTCCATTGACCGCAGTCC	Sox3	X94125]
st97_793	CATGGTCTGGTCGCAAATC	CCATGTGCTTTAGCCGAAGT	Sox19	X98368
st97_794	GAGCGGAGAAAAATCATGGA	AGTCCGCCATGTGCTTGAG	Sox12	Z18961

2.4.1 Results of the PCR for the Sox Genes

The tissue profiles of the PCR products, together with a summary of the PCR conditions, for each of the Sox gene primer sets used are shown as Figures 36-62. In order to view these figures graphically, a number of the original images were digitized as described in Chapter 1, section 1.5. Figure 35 is a list of the tissues examined with identification numbers for samples to accompany the graph images of tissue profiles for the *Sox* genes 1, 2, 4, 9 and 15. Each of the *Sox* genes will be discussed in the light of the expression profiles and existing knowledge in the published literature.

Legend for all gel images Figures 36 – 62 is:

Ethidium bromide stained 2.5% agarose gels, run at 125 volts, 30 milliamps, 2 hours and photographed with a Polaroid camera. Photographs were scanned on an Epson scanner, opened in Adobe Photoshop, where defined regions were then cut and pasted into MacDraw and scored through viewing the original photographs.

Gene/Est: Gene or EST name
Symbol: Acronym for gene/est
EMBL AcNo: Accession number from EMBL database
Primers: Sanger Centre identifying number

PCR: Polymerase Chain Reaction
Date: date of lab. work
$MgCl_2$: final concentration
Temp: annealing temperature
Cycles: number of cycles

Amplification signal: four catagories
Strong – dark blue – major signal from both duplicates.
Moderate-weak – blue – broadest range of signal.
Trace – green – faint signal.
None – yellow – no signal from either duplicate.
Discrepancy – red – signal from one only, of the duplicates.

Figure 35: Tissue identification of graph position

ID	tissue
1	whole conceptuses 8.5d
2	whole conceptuses 9.5d
3	whole embryo 10.5d
4	whole embryo 11.5d
5	whole embryo 12.5d
6	whole embryo 13.5d
7	whole embryo 15.5d
8	whole embryo 17.5d
9	forebrain 12.5d
10	forebrain 13.5d
11	forebrain 15.5d
12	forebrain 17.5d
13	midbrain 12.5d
14	midbrain 13.5d
15	midbrain 15.5d
16	midbrain 17.5d
18	hindbrain 12.5d
19	hindbrain 13.5d
20	hindbrain 15.5d
21	hindbrain 17.5d
22	spinal cord 12.5d
23	spinal cord 13.5d
24	spinal cord 15.5d
25	spinal cord 17.5d
26	head 12.5d
27	head 13.5d
28	head 15.5d
29	head 17.5d
30	heart 12.5d
31	heart 13.5d
32	heart 15.5d
33	heart 17.5d
35	lung 12.5d
36	lung 13.5d
37	lung 15.5d
38	lung 17.5d
39	liver 12.5d
40	liver 13.5d
41	liver 15.5d
42	liver 17.5d
43	gut 12.5d
44	intestine 13.5d
45	oes/stom 13.5d
46	intestine 15.5d
47	oes/stom 15.5d
48	intestine 17.5d
49	urogenital/kidney 12.5d
50	urogenital /kidney 13.5d

ID	tissue
52	urogenital/gonads 15.5d
53	urogenital male gonads 17.5d
54	urogenital female 17.5d
55	bladder 15.5d
56	kidney 15.5d
57	bladder 17.5d
58	kidney 17.5d
59	spleen 17.5d
60	forelimbs 12.5d
61	forelimbs 13.5d
62	forelimbs 15.5d
63	forelimbs 17.5d
64	rest of body 12.5d
65	rest of body 13.5d
66	rest of body 15.5d
67	rest of body 17.5d
69	extra embryonic 10.5d
70	extra embryonic 11.5d
71	placenta 12.5d
72	yolk sac 12.5d
73	placenta 13.5d
74	yolk sac 13.5d
75	placenta 15.5d
76	yolk sac 15.5d
77	placenta 17.5d
78	yolk sac 17.5d
79	whole brain - adult
80	spinal cord - adult
81	skeletal muscle - adult
82	heart - adult
83	liver - adult
84	kidney - adult
86	fundus - adult
87	caecum - adult
88	testis - adult
89	ovary - adult
90	one day old mouse
91	glycogen
92	mouse genomic
93	human genomic
94	rat genomic
90	glycogen

Figure 36: Expression Profile of *Sox1* in Mouse Foetal Panel

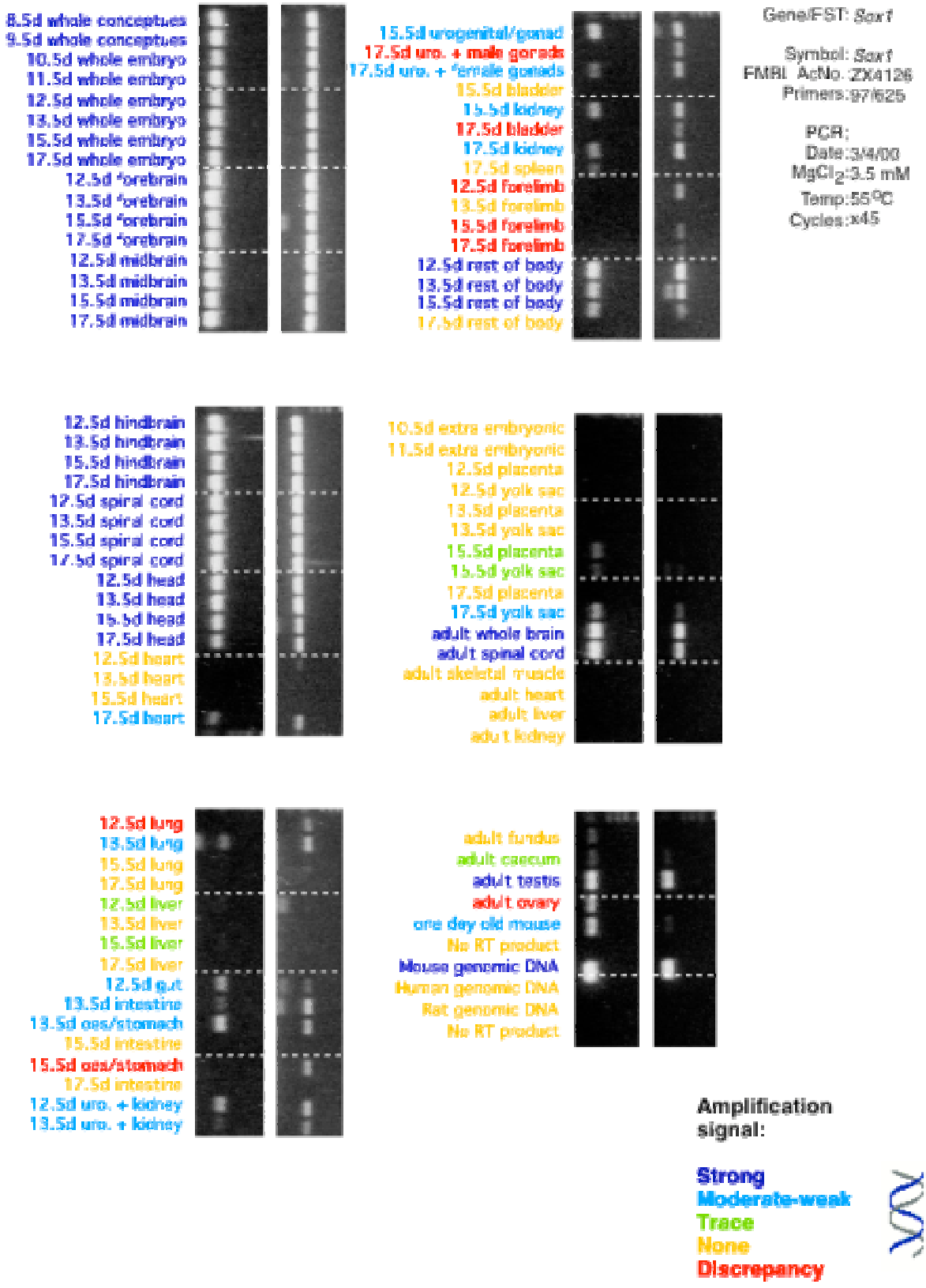
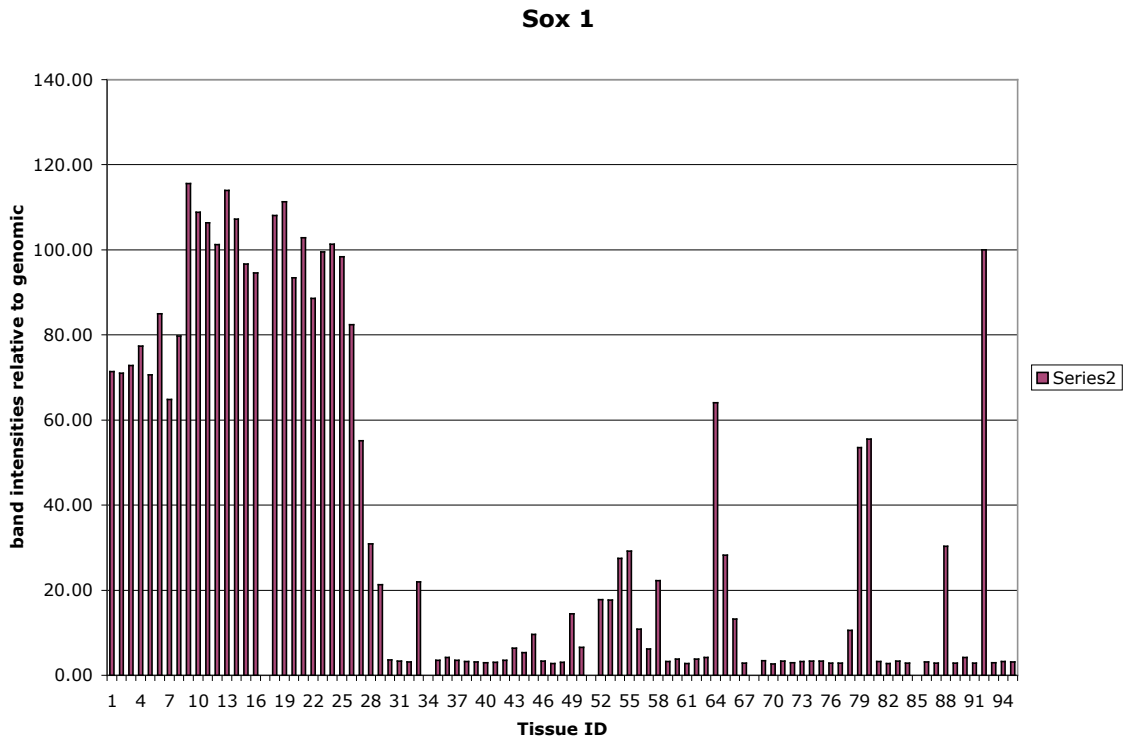


Figure 37: Graphical representation of gel image in Figure 36.



Densitometric measurements of gel band fluorescence displayed as a percentage of the mouse genomic fluorescent band intensity in position 92. Tissue identifications are located in Figure 35.

Sox1

Figure 36: *Sox 1* is noticeably strong in the whole organism, most probably as a result of the strong gene expression found in the brain and spinal cord. Internal organs show a much lower expression.

The graph Figure 37, of *Sox 1* highlights the overall pattern, clearly illustrating highest expression in the whole embryo, brain, spinal cord and head (1-29), plus rest of body13.5d – 15.5d (64-66); with strong expression also in the adult tissues brain (79), spinal cord (80), testis (88). Expression also noted in heart 17.5d (33) – but not in heart for 12.3d to 15.5d; and some of the internal organs - urogenital 15.5d, 17.5d, plus bladder 15.5d (52-55), and spleen 17.5d (58).

The adult tissues mirror expression found [2] earlier and confirm the importance of *Sox 1* in the brain, spinal cord and testis. Early studies have implicated *Sox 1*as an important gene in mouse brain [3], eye development [4] and it has also been shown to be involved in the onset of mouse neural determination [5].

Sox2

Figure 38: *Sox 2* expression in this study is shown through out the developing foetus with significantly lower levels in the liver as shown by the graph of the gel band intensities in Figure 38, and extra embryonic regions of early stages plus placenta and yolk sac of the later stages. The adult stages show no expression in the liver or kidney with strong expression in the neuronal tissues and fundus.

From the *Sox 2* graph, Figure 39, tissues with low expression are clearly identifiable as 12.5d heart (30), 15.5d heart (32), liver 12.5d – 17.5d (38-42), 17.5d intestine (48), and, although not obvious from the gel, spleen 17.5d (59).

The adult pattern of expression confirms earlier studies [2]. In mice, maternal *Sox2* is operational until the late morula stage, when zygotically transcribed *Sox 2* becomes essential for normal epiblast growth at post implantation stages [6]. Knockout *Sox2* mice are embryonic lethals. *Sox 2* (together with *Pax6*) are required for murine lens differentiation [7], [8], [9]. In other species, *Sox 2* is implicated in ovine developing gonads and germ cell formation [10].

In *Drosophila*, Dichaete, a protein related to *Sox 2*, is found throughout the developing CNS (central nervous system) [11], and from studies in chicken, *cSox2* expression increases in the CNS as neural ectoderm is established [12].

Figure 38: Expression Profile of *Sox2* in Mouse Foetal Panel

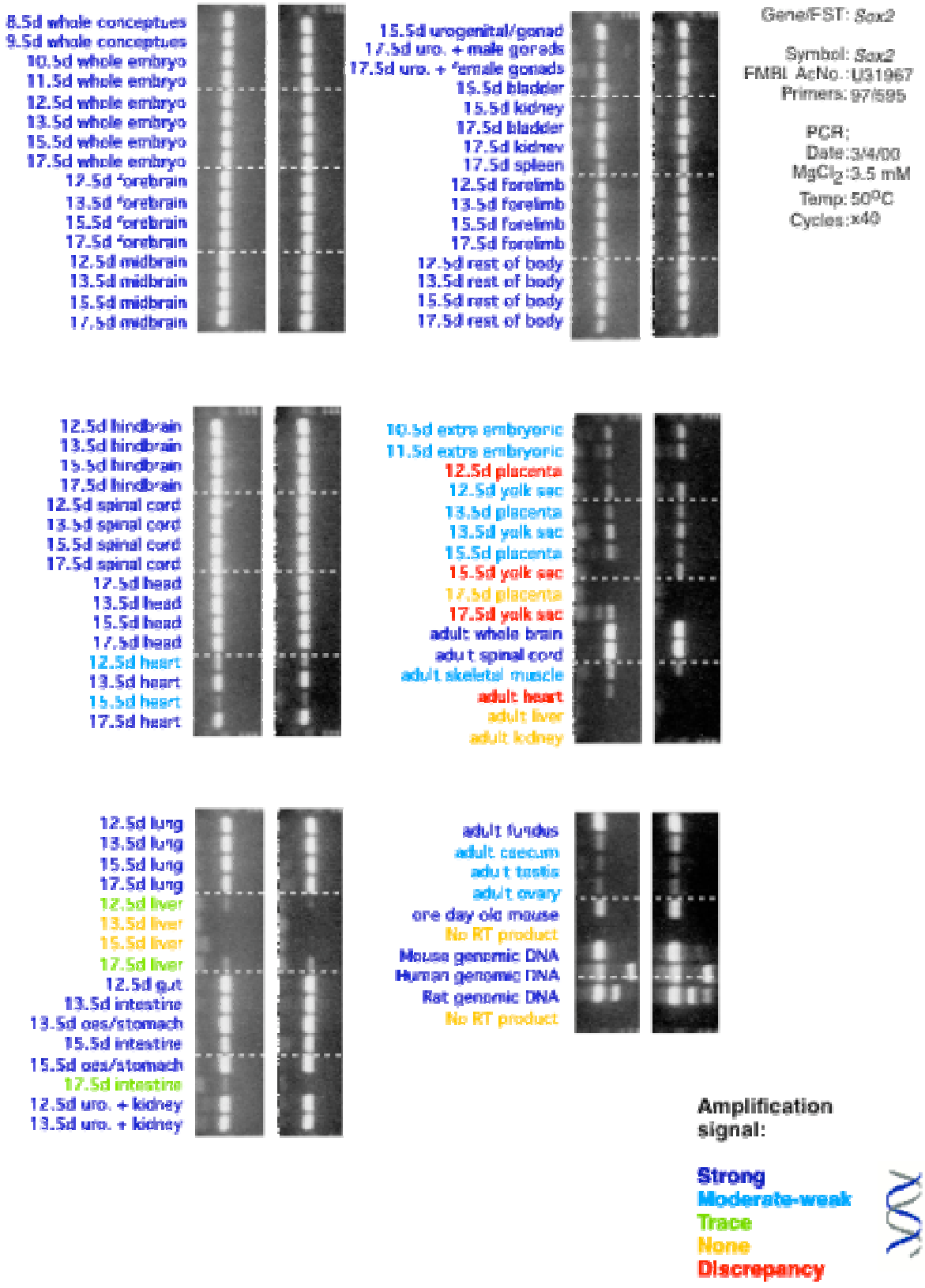
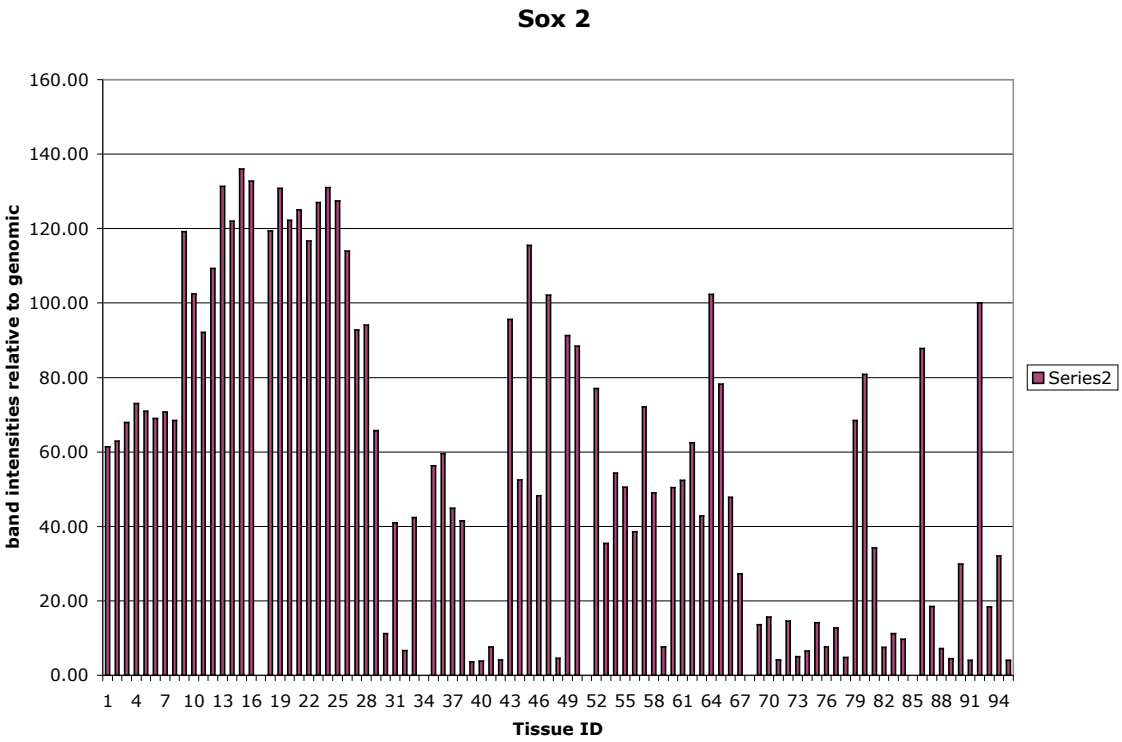


Figure 39: Graphical representation of gel image in Figure 38



Densitometric measurements of gel band fluorescence displayed as a percentage of the mouse genomic fluorescent band intensity in position 92. Tissue identifications are located in Figure 35.

Sox4

Figure 40: *Sox 4* expression is found throughout the developing and adult mouse in this study and previously [2]. The graph, Figure 41 of the gels for *Sox 4* highlights areas with lower levels of expression as head 17.5d (29), heart 15.5d (32) and liver tissues (39-42). *Sox4* has been reported in a wide range of tissues and knockout mice have been shown to die of heart defects during embryogenesis with malformed CNS [13].

Figure 40: Expression Profile of *Sox4* in Mouse Foetal Panel

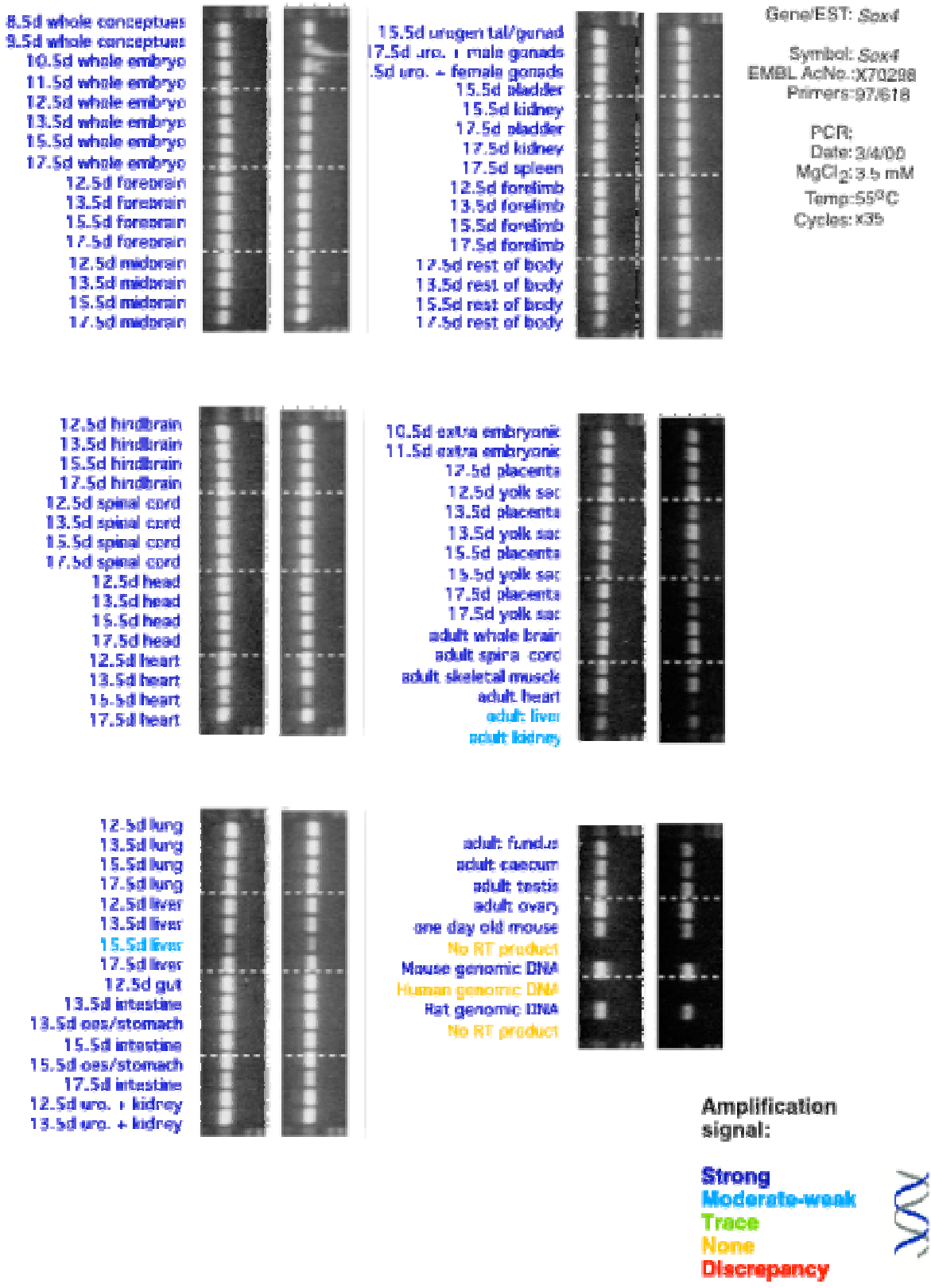
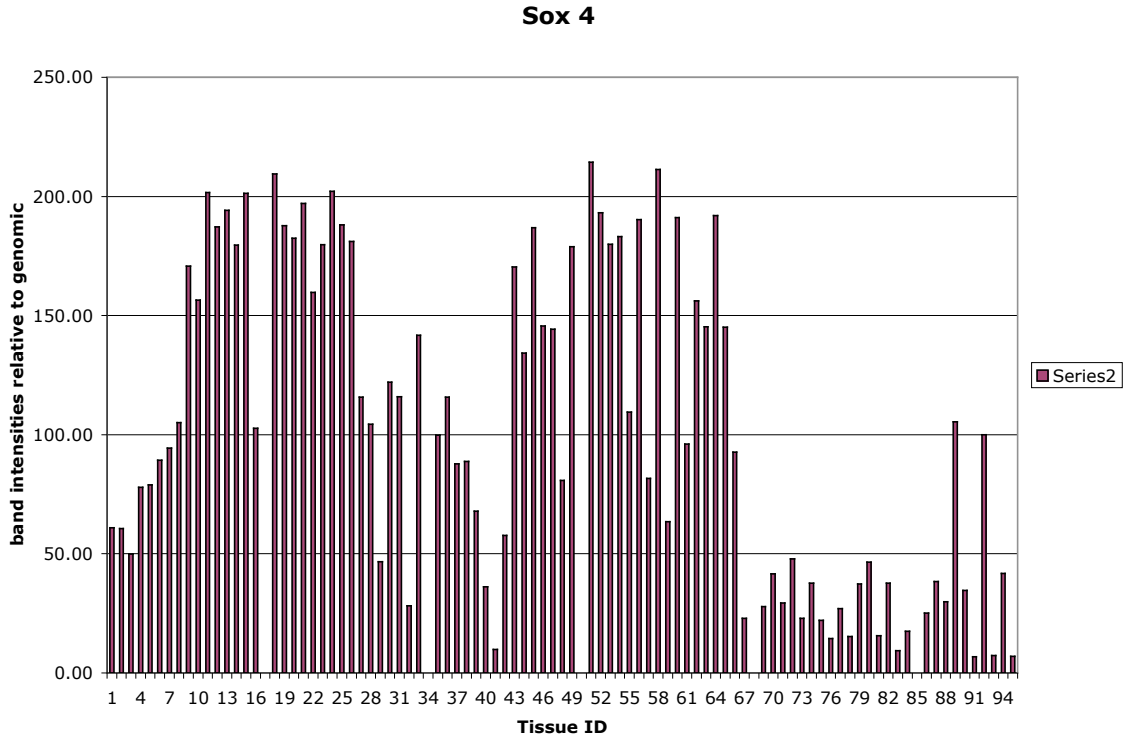


Figure 41: Graphical representation of gel image in Figure 40.



Densitometric measurements of gel band fluorescence displayed as a percentage of the mouse genomic fluorescent band intensity in position 92. Tissue identifications are located in Figure 35.

Sox5

Figure 42: *Sox 5* expression in the mouse foetal panel is clearly greater than that found in the adult panel [2]. *Sox 5* has been reported to be important during chondrogenesis [14] [15], cartilage formation being a major activity during development, and this would fit with the high levels found throughout this foetal panel. Post-meiotic spermatids in adult testis [16] have also been identified as a major site for *Sox5* expression, hence the large amount found here in adult testis. It has been reported that the human form of *SOX5* exists as different transcripts in the fetal brain and adult testis [17], which may explain the difference in levels between adult and foetal development found in this study.

Figure 42: Expression Profile of *Sox5* in Mouse Foetal Panel

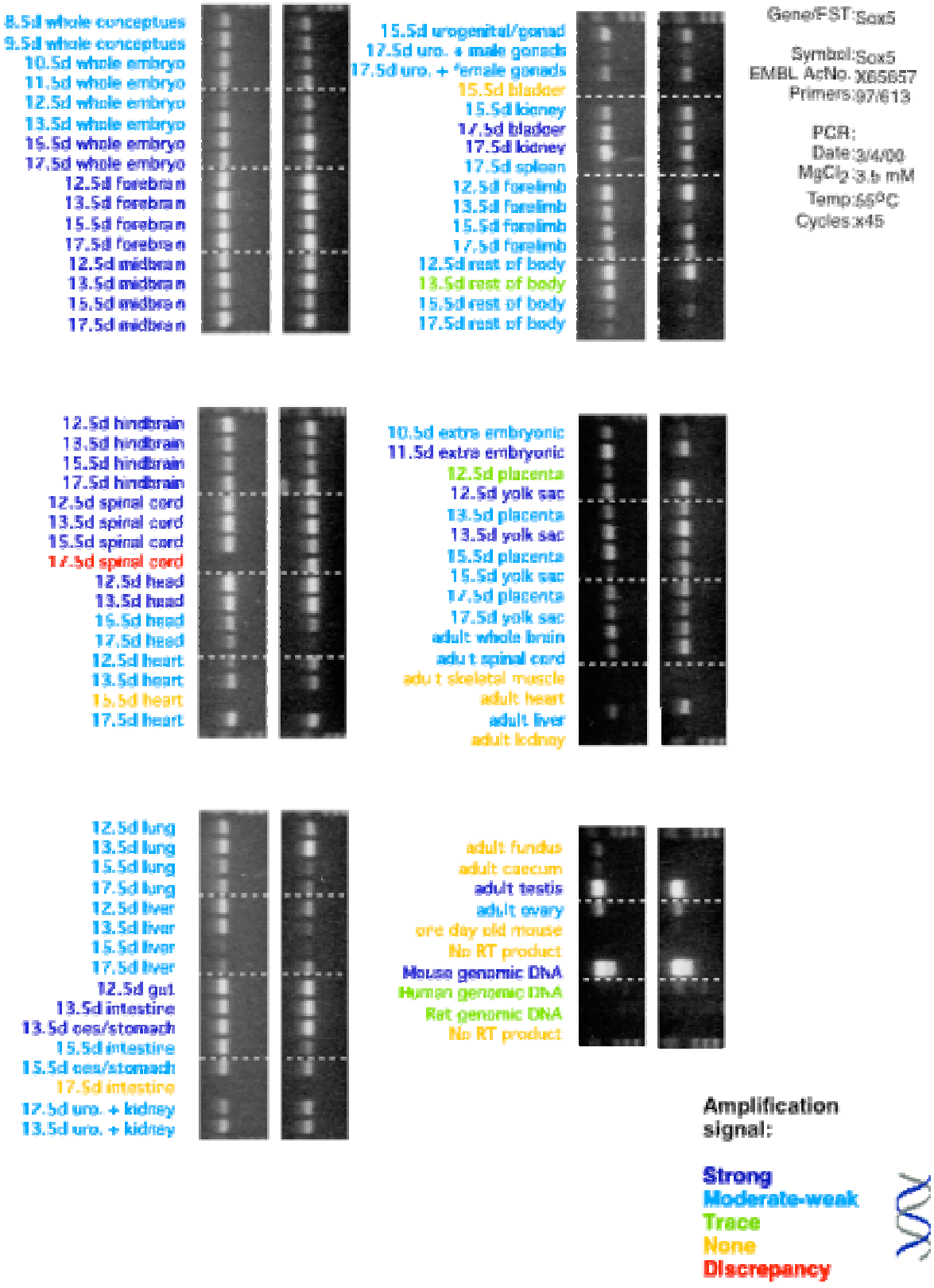
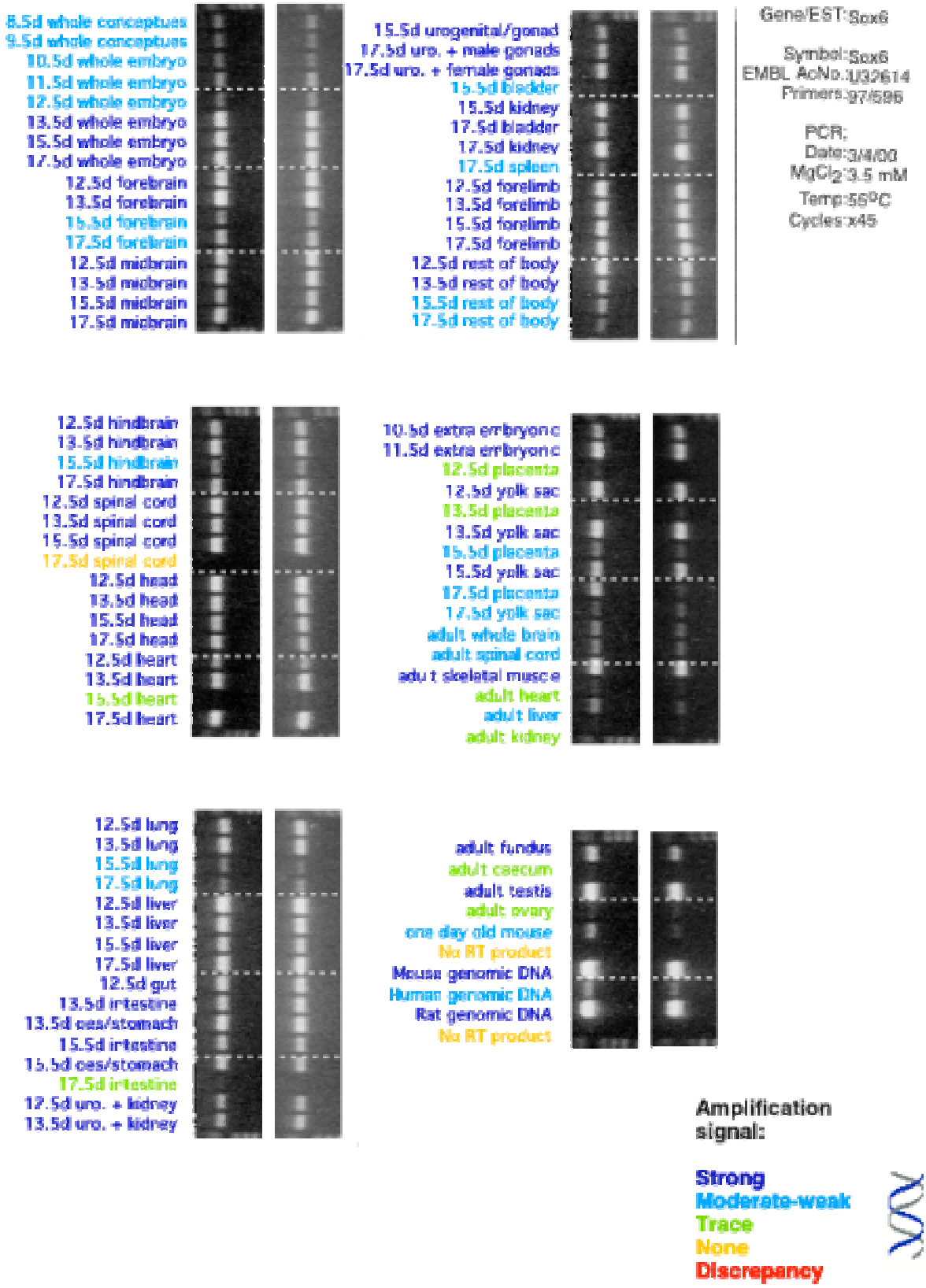


Figure 43: Expression Profile of *Sox6* in Mouse Foetal Panel



Sox6

Figure 43: *Sox6* pattern of expression in this foetal panel is not dissimilar to that found for *Sox5* (Figure 42), with the major exception of the adult skeletal muscle (*Sox6,* but not *Sox5)*. *Sox5* and *Sox6* are very similar at the sequence level, clustering in the B group [18]. Like *Sox5, Sox6* is required for chondrogenesis [14] and would explain the similarity in expression found in this study.

Sox7

Figure 44: *Sox7* shows a wider expression profile in this study relative to the expression profile found earlier [2] in adult tissues alone. Northern analysis of three adult tissues found ovary and heart positive but not skeletal muscle [19], confirming the expression pattern found in the adult tissues by RTPCR. From *in-situ* hybridization and Northern blot analyses, it has now been shown to be present in a variety of adult and fetal tissues, including skeletal muscle and is now thought to be involved in the development of the vascular system [20].

Sox8

Figure 45: *Sox8* expression in this study is found throughout the developing foetus in all tissues, with significantly less in both 15.5d and 17.5d lung and liver. Studies in *Sox8*-deficient mice, show this gene is not essential for development, suggesting that it's role may be taken by one of the other group E family members, either *Sox9* or *Sox10* [21]. It has also been suggested that this gene may substitute for *Sox9* during the formation of Sertoli cells. *Sox8* has been shown to be expressed between 12dpc in mouse testis and beyond 16dpc in the Sertoli cells [22]. From

work with cartilage development, all three E group *Sox8, 9* and *10* genes are thought to operate in a concerted fashion to effect digit cartilage [15].

Figure 44: Expression Profile of *Sox7* in Mouse Foetal Panel

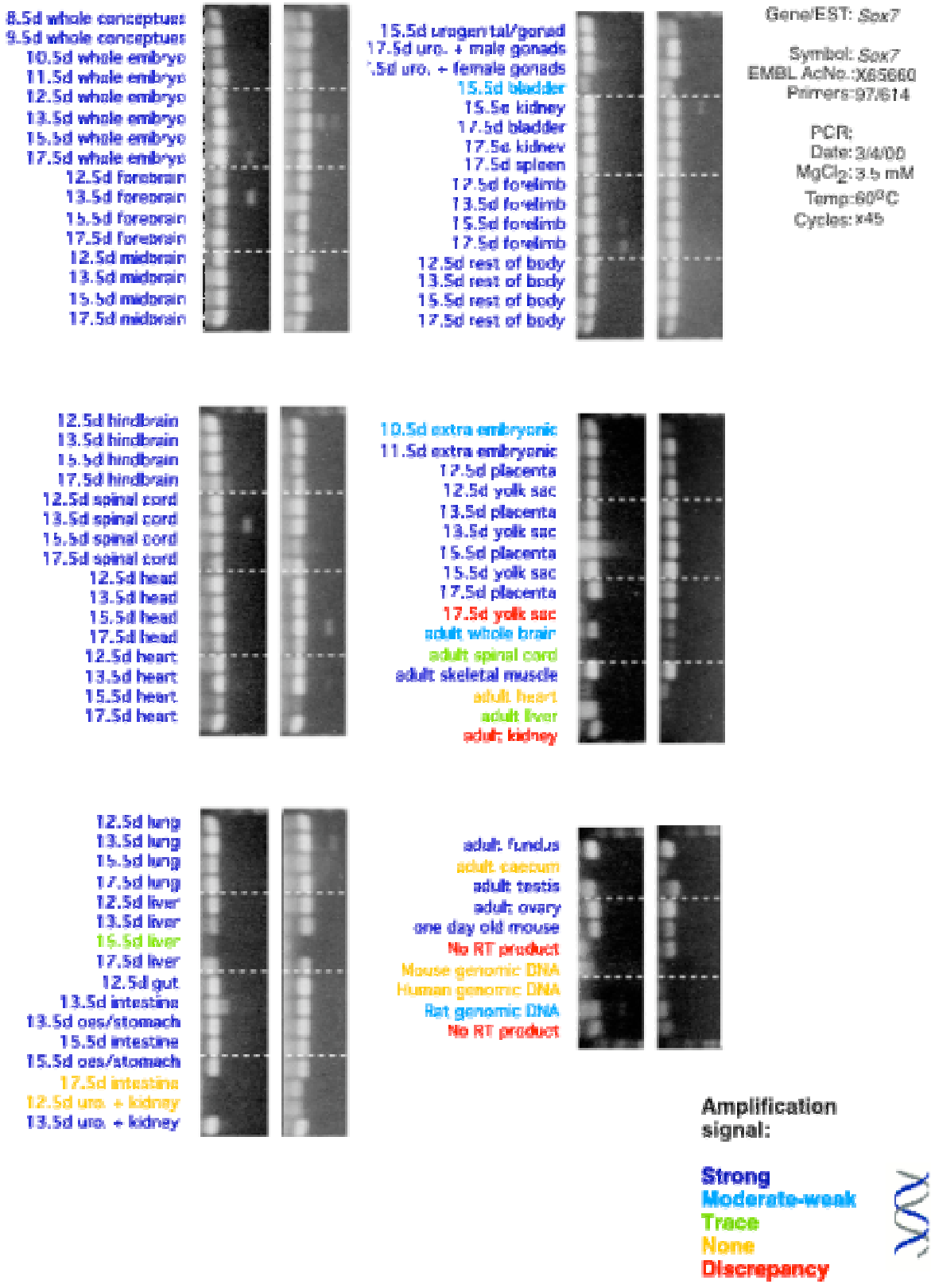


Figure 45: Expression Profile of *Sox8* in Mouse Foetal Panel

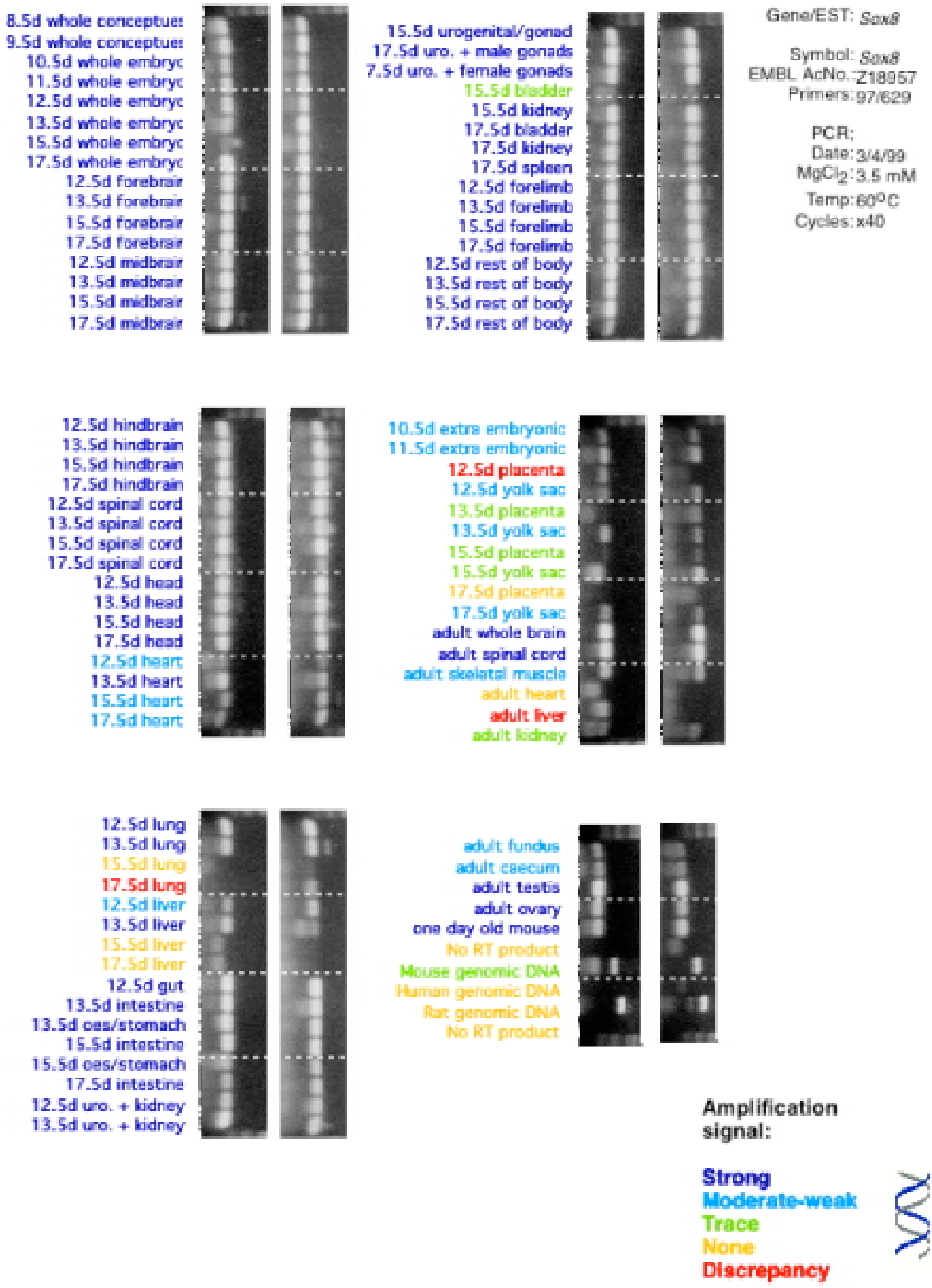


Figure 46: Expression Profile of *Sox9* in Mouse Foetal Panel

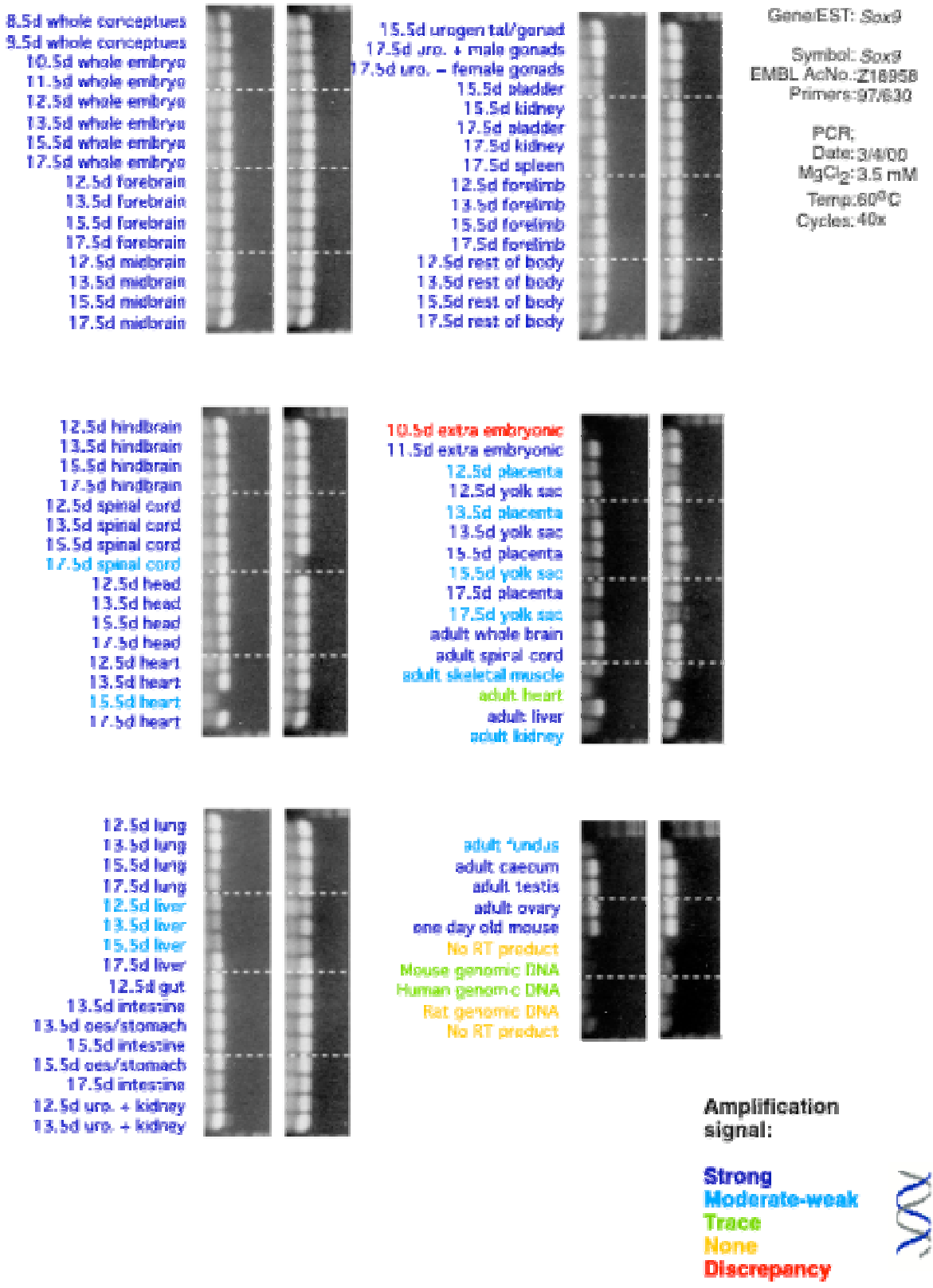
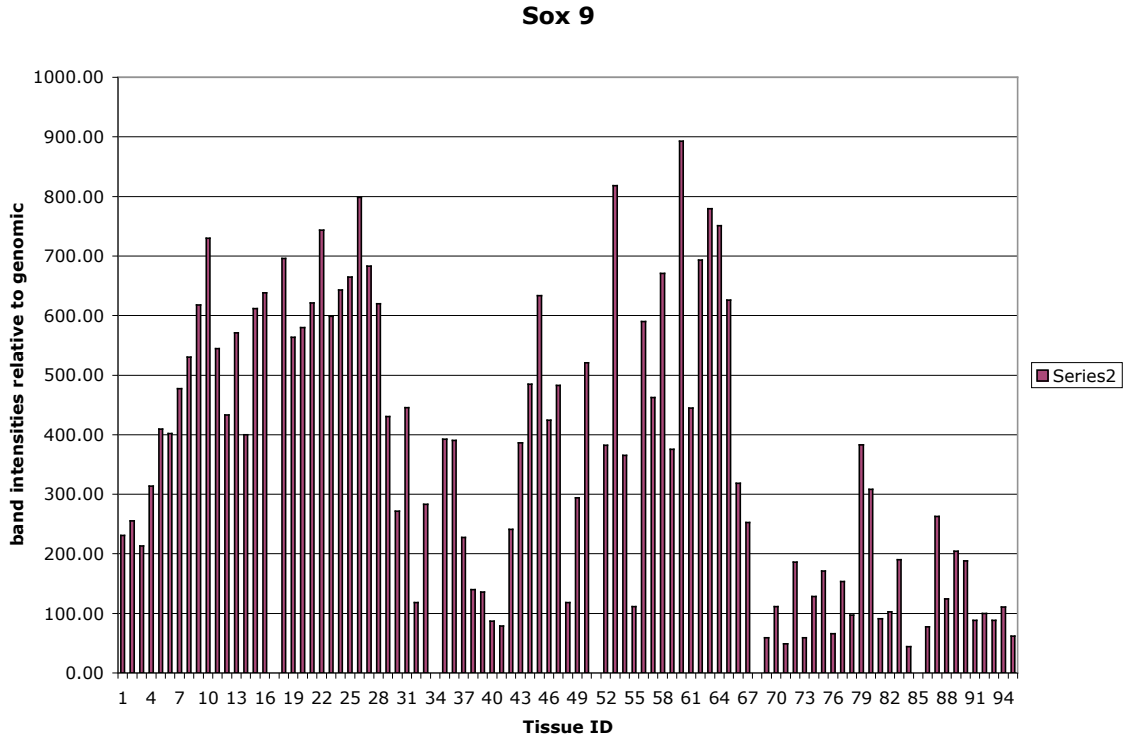


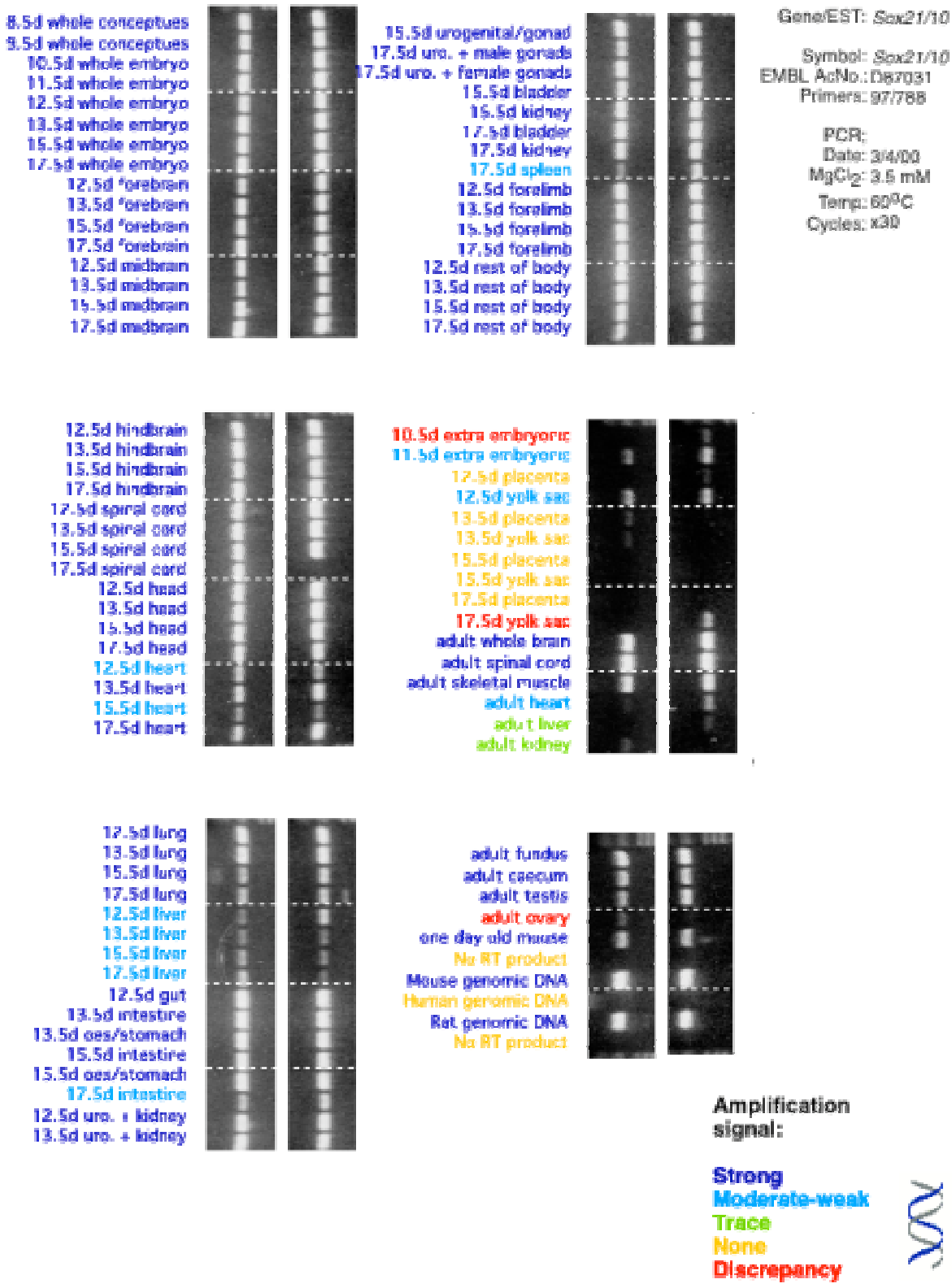
Figure 47: Graphical representation of gel image in Figure 46.



Sox9

Figure 46: *Sox9* expression is found in all tissues of the developing and adult mouse, in this and earlier studies [2]. The graph, Figure 47, of this gel highlights the regions of highest expression as brain (9-21), spinal cord (22-25), regions of the oesophagus/stomach/intestine (43-49), plus lower expression found in liver samples (39-41) 12.5d, 13.5d and 15.5d. All the urogenital regions except for 15.5d bladder show high expression, as does the placental and yolk sac tissues. From the adult section, brain (79), spinal cord (80) and caecum (87) are showing greatest expression. A lack of this gene leads to the disease campomelic dysplasia, a bone malformation and sex reversal [23]. Sox 9 is important in the maintenance of Sertoli cell development during testis formation [24] and implicated as the main element driving cartilage formation through binding to the type XI collagen gene Col11a2 [25].

Figure 48: Expression Profile of *Sox10* in Mouse Foetal Panel



Sox10

Figure 48: *Sox10* and *Sox21* were initially thought to be two different sequences, but they have since been assigned the same name of *Sox10* [26]. *Sox10* in this study is found throughout the developing foetus with the exception of placenta and from 13.5d the yolk sac. It is also low in all the liver tissues. Mutations in *Sox10* cause a colon malformation, identified in humans as Waardenburg-Hirschsprung disease [27]. *Sox10* has been found associated with developing and mature glial cells of the CNS, and Northern blot and *in-situ* analysis has shown the *Sox10* transcript in the brain regions of mouse (and rat) [28]. Western analysis identified *Sox10* protein in cultured rat oligodendrocytes and Schwann cells. Whole mount *in-situ* illustrated the presence of *Sox10* in a wide variety of neural tissue, including facial-cranial ganglia, brain regions, spinal cord and outer wall of stomach [28]. It may be of interest to note the absence of expression in maternal tissues for 13.5d – 17.5d, given the widely found expression in foetal and adult tissues.

Sox11

Figure 49: *Sox11* in this study is found in the full range of developing tissues, with lower levels found in liver samples. The adult portion of this panel reflects expression patterns found in an earlier study [2], confirming the high levels found in neuronal tissue and testis. From studies in chick, *cSox11* has been shown to be associated with the neural epithelium [29] and the developing CNS and PNS (peripheral nervous system) of mice [13].

Figure 49: Expression Profile of *Sox11* in Mouse Foetal Panel

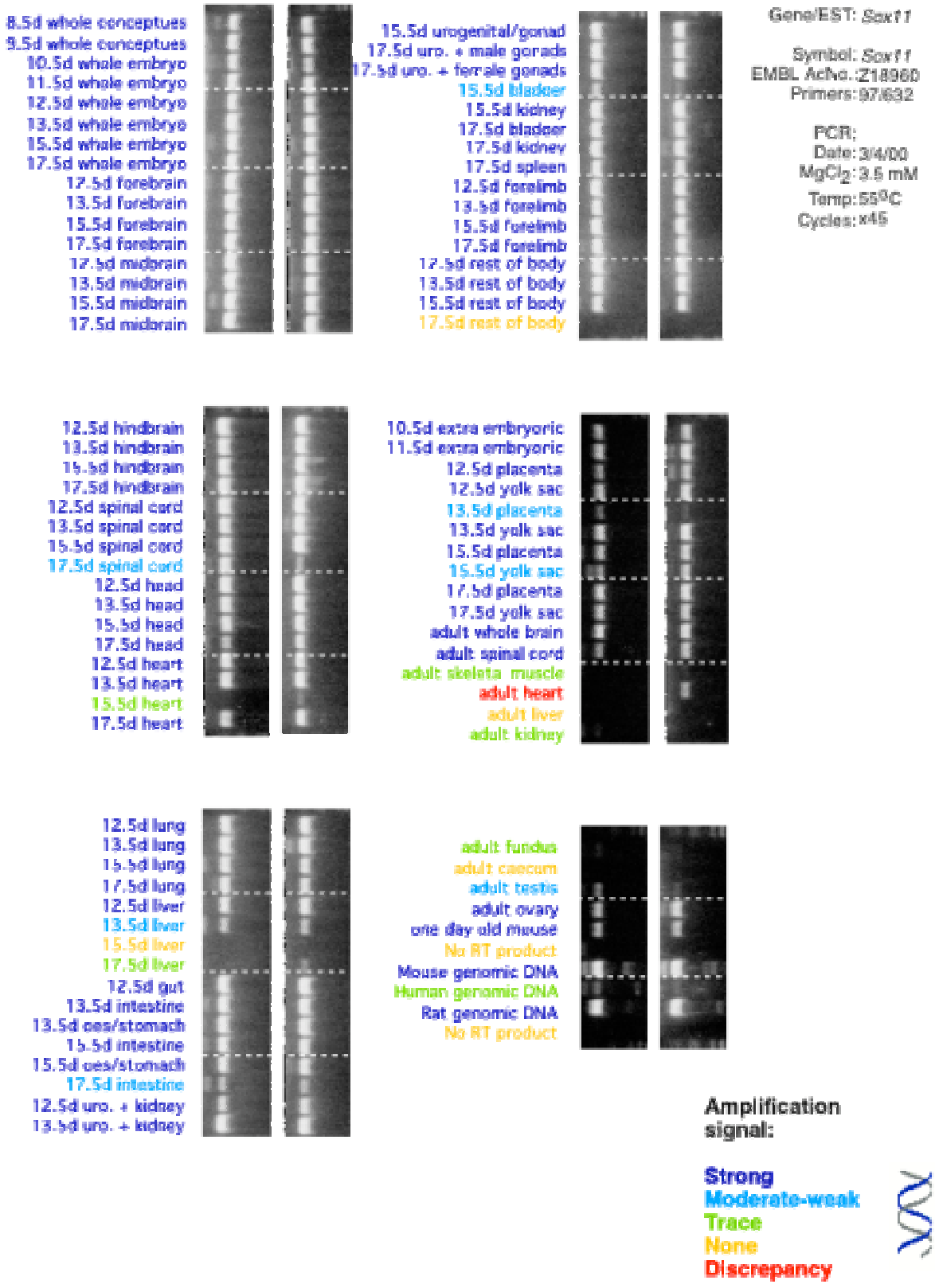
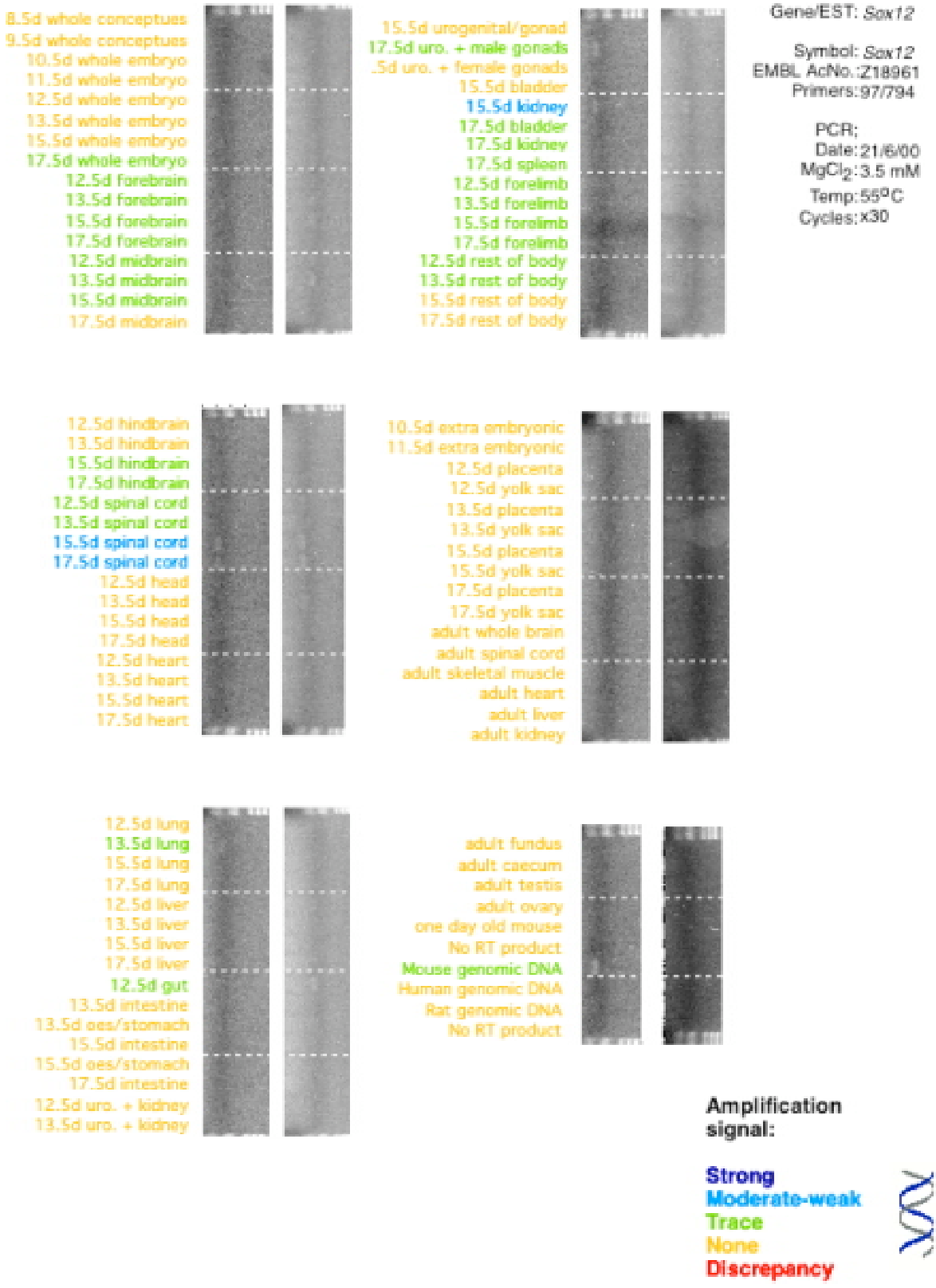


Figure 50: Expression Profile of *Sox12* in Mouse Foetal Panel



Sox12

Figure 50: *Sox12* in this study showed faint expression in some brain regions, spinal cord, 13.5d lung, 12.5d gut 17.5d urogenital male tissues and 15.5d kidney. There is very little reported on this gene since it was initially discovered [30]. However, *Sox22* has been renamed *Sox12* [31] and studies on the human *SOX22* show expression in a variety of adult and fetal tissues and organs, so it is plausible to expect *Sox12* to be similarly important in the differentiation and maintenance of a range of cell types [32].

Sox13

Figure 51: *Sox13* expression in this study shows expression in all tissues throughout the adult and foetus. An earlier report of Northern analysis showed the *Sox13* transcript to be restricted to ovary and kidney [33] (other tissues tested include brain, thymus, spleen and testis). This group D Sox gene expression is reported to be associated with the development of arterial walls [31].

Sox14

Figure 52: *Sox14* showed a very distinct pattern of expression in this study of foetal tissues. Undetected in very early stages (8.5d, 9.5d and 10.5d), with high expression in all brain and spinal cord regions through out development to adult. Expression was also found at 15.5d for intestine and oesophagus/stomach, urogenital/gonad and yolk sac; 17.5d urogenital male specific gonad and yolk sac; rest of body for 12.5d and 13.5d, with trace amounts at 15.5d. The adult section of this panel reflects that found in our earlier study [2]. The *Sox14* expression has been found in a highly restricted pattern in neurons in the developing mouse brain and

spinal cord [34], confirming the findings in this study.

Figure 51: Expression Profile of *Sox13* in Mouse Foetal Panel

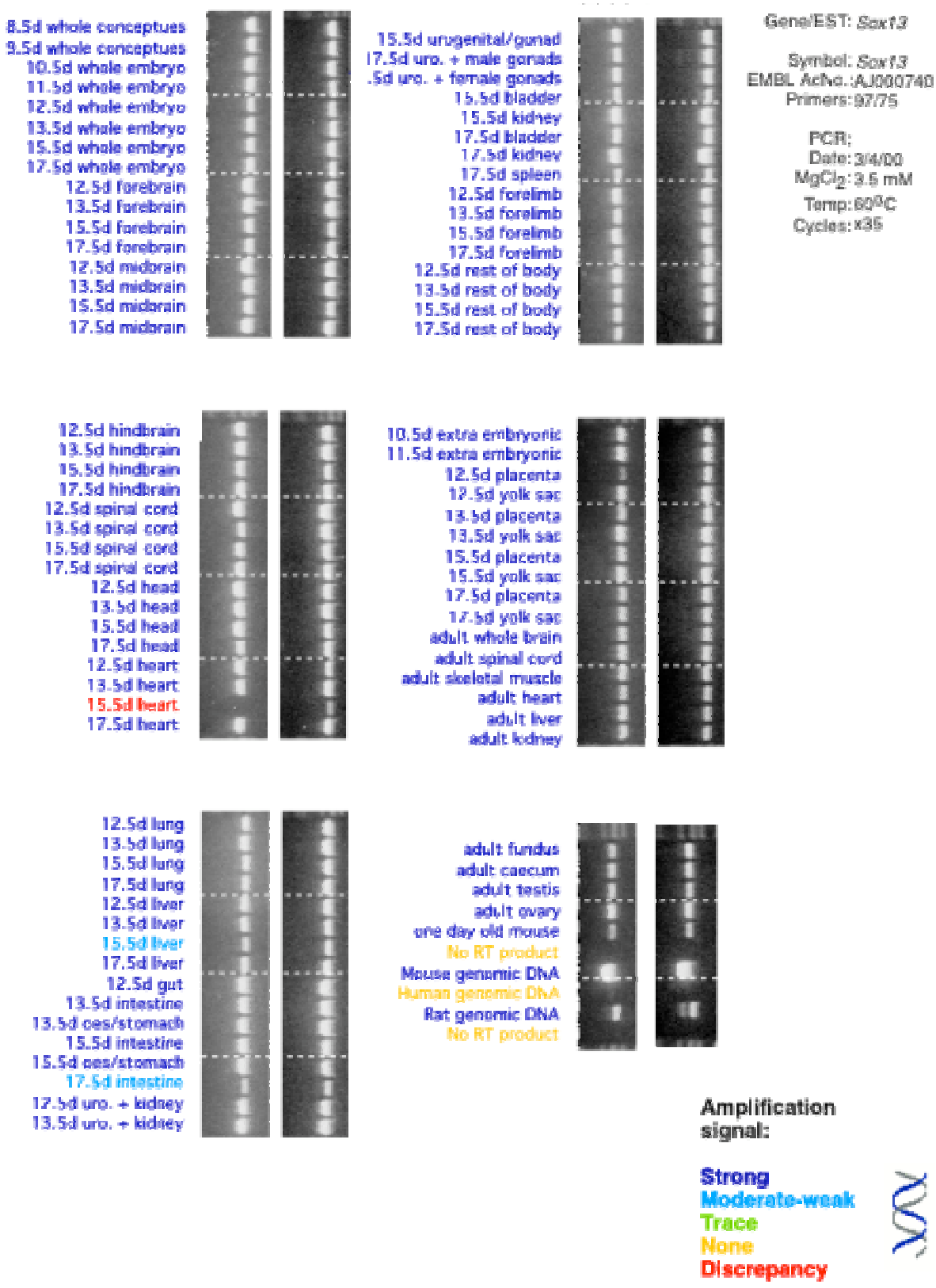


Figure 52: Expression Profile of *Sox14* in Mouse Foetal Panel

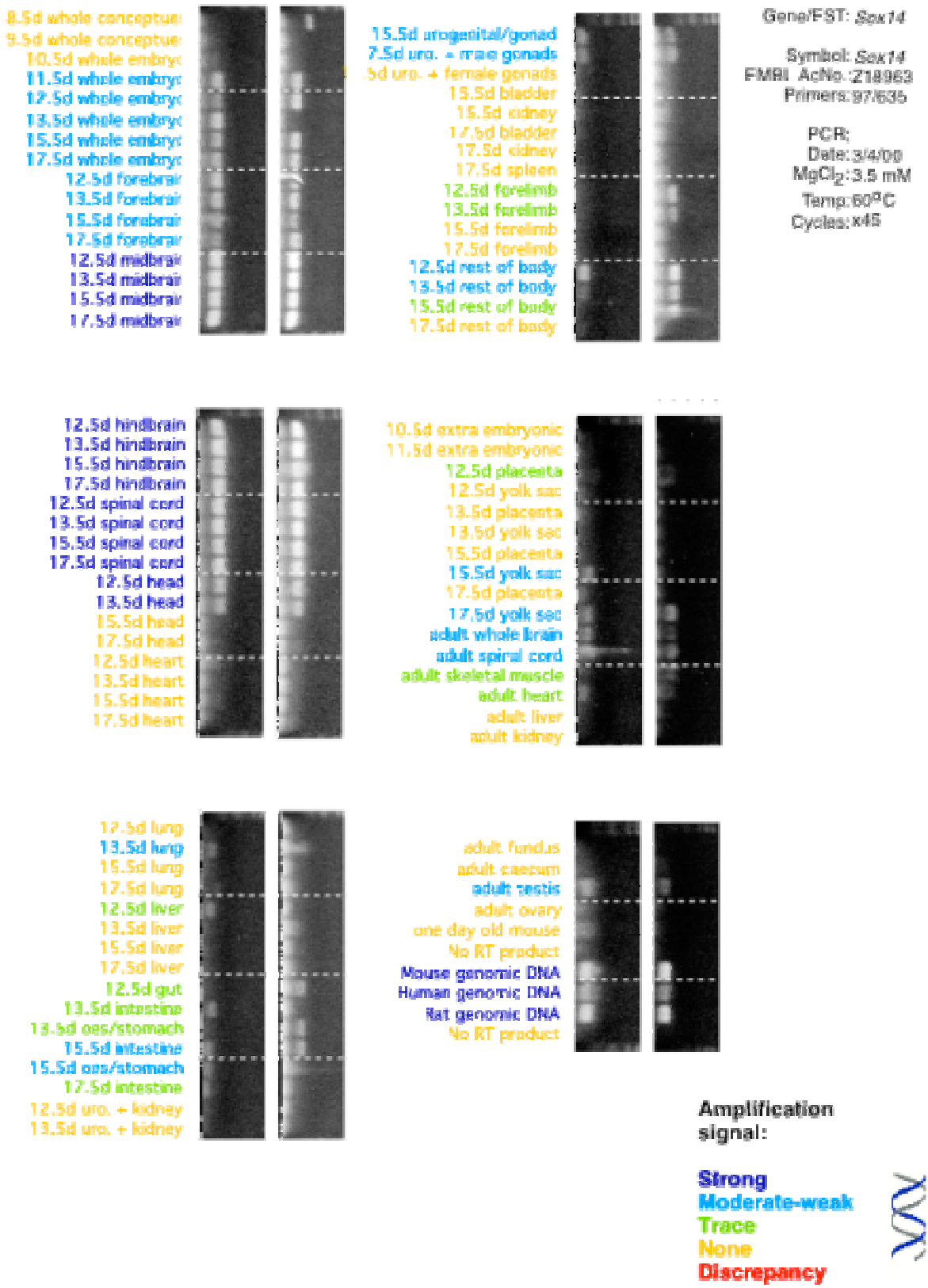


Figure 53: Expression Profile of *Sox15* in Mouse Foetal Panel

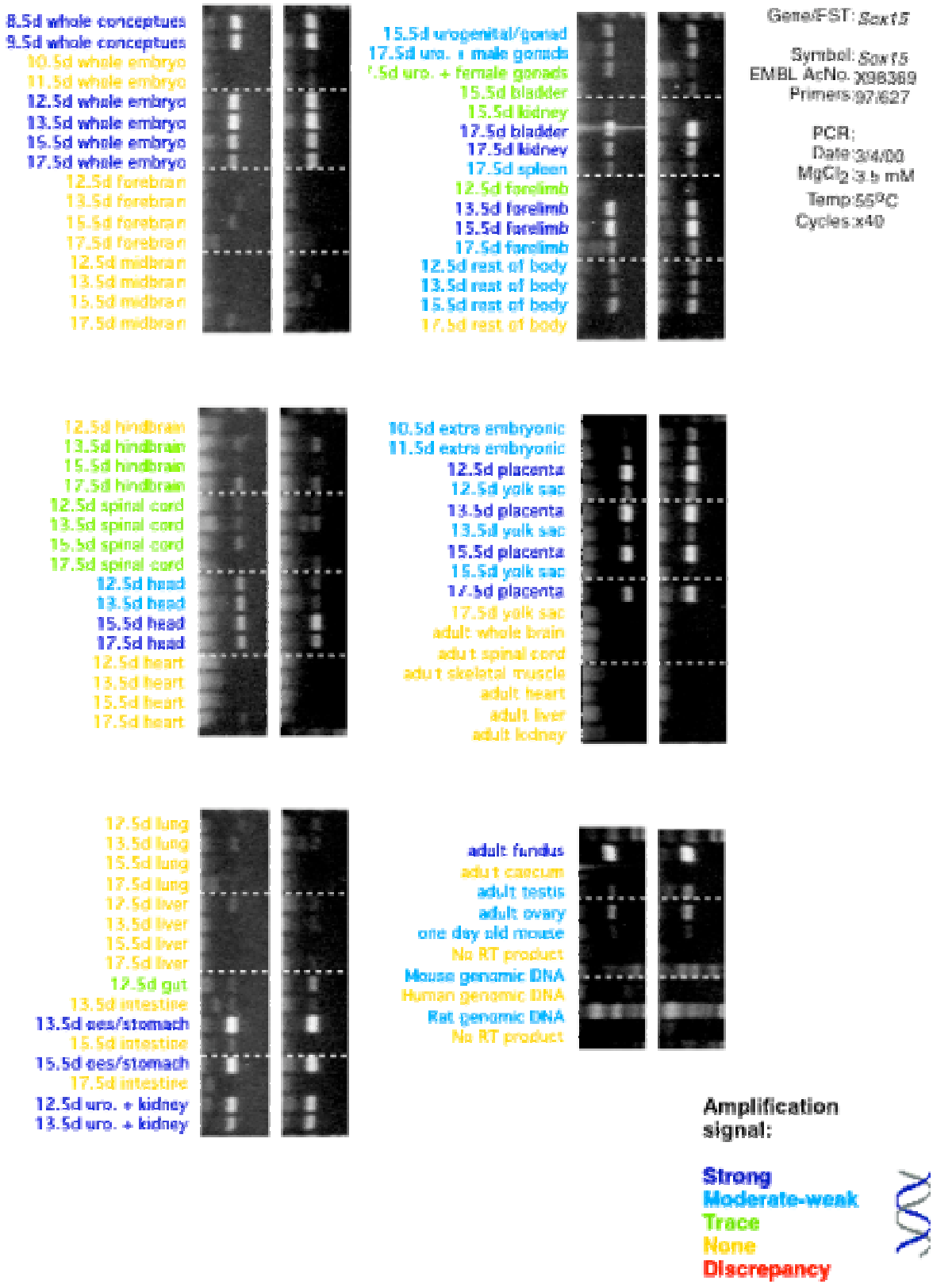
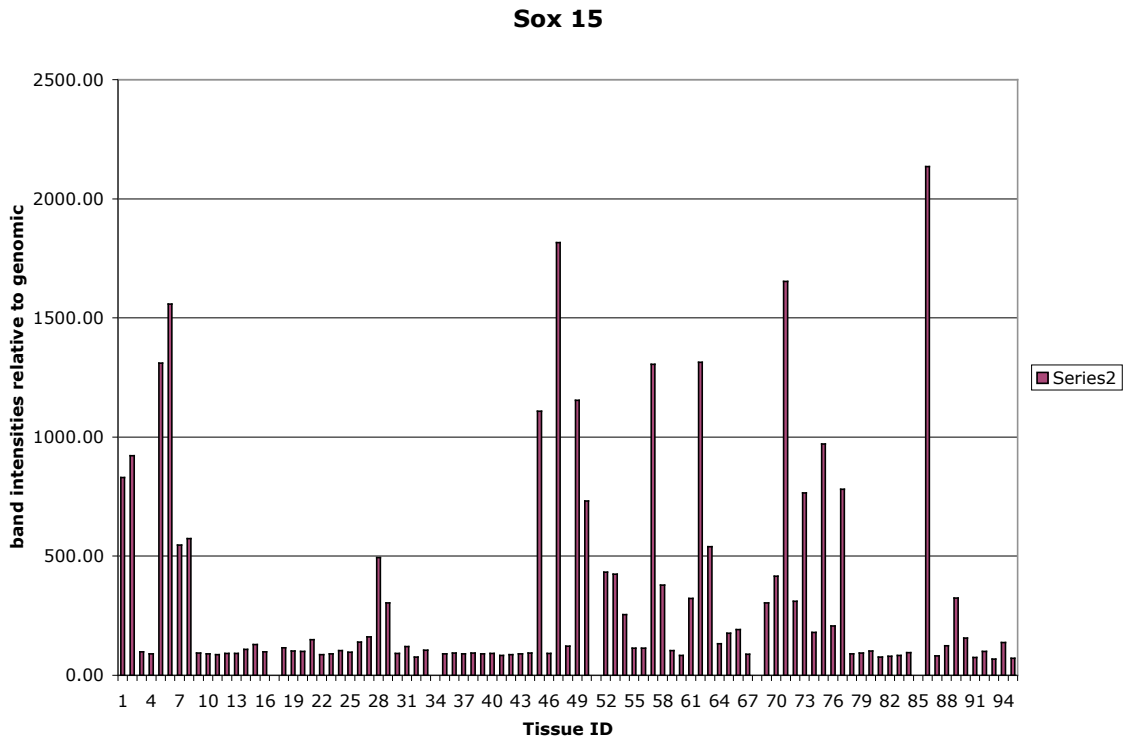


Figure 54: Graphical representation of gel image in Figure 53.



Densitometric measurements of gel band fluorescence displayed as a percentage of the mouse genomic fluorescent band intensity in position 92. Tissue identifications are located in Figure 35.

Sox15

Figure 53: *Sox15* expression in this study illustrates high levels of *Sox15* in a variety of tissues. The whole organism for 10.5d and 11.5d interestingly gave no expression. Trace amounts of expression were found in the hindbrain and spinal cord, with no expression in forebrain and midbrain, though expression was found in the remainder of the head. No expression was detected in heart, lung, liver or intestine, with high amounts in oesophagus and stomach, urogenital regions of 12.5d, 13.5d, bladder 17.5d, kidney 17.5d. More expression was found in 13.5d and 15.5d forelimb relative to 12.5d and 17.5d, but less in the rest of body at 17.5d relative to earlier stages, and more in placenta relative to yolk sac. Figure 54 is the graphical

representation of the gel, illustrating minimal expression from head regions (28 & 29), significant expression from oesophagus/stomach regions (45, 47) and urogential/kidney (49 & 50) relative to intestine and is mirrored in the adult panel where fundus (85) has much higher expression than caecum (86). The adult section relates to the expression pattern found earlier [2].

It has been reported that the human *SOX15* is expressed in skeletal muscle and is required for myogenesis [35]. However, this report is not confirmed in this study on mouse.

Sox16

Figure 55: *Sox16*, like *Sox15,* showed a variety of expression patterns in these tissues. Undetected in the 11.5d whole organism, with expression found in all other stages for whole organism. Brain and spinal cord regions showed very low or no levels of expression, but again expression was detected in the remainder of the head. Heart, lung and liver tissues showed only trace expression in 17.5d lung and 12.5d liver. Expression was found in the gut regions of 12.5d to 15.5d (though not intestine tissues) and urogenital (urogenital/gonad) regions for all stages. It was detected in the 17.5d bladder and kidney, but undetected in these tissues for 15.5d. The pattern of forelimb expression found for *Sox15* was mirrored here, with more found in 13.5d and 15.5d than either 12.5d or 17.5d. Also 17.5d rest of body showed no expression. Again, mirroring the *Sox15,* placental tissues showed a higher expression relative to the yolk sac for all stages. In adult tissues, *Sox16* was detected in skeletal muscle, fundus and ovary, with trace amounts in spinal cord and testis.

Little is known about the *Sox16* gene, though it is reported to have strong homologies with *Sox15* at the sequence level [36].

Figure 55: Expression Profile of *Sox16* in Mouse Foetal Panel

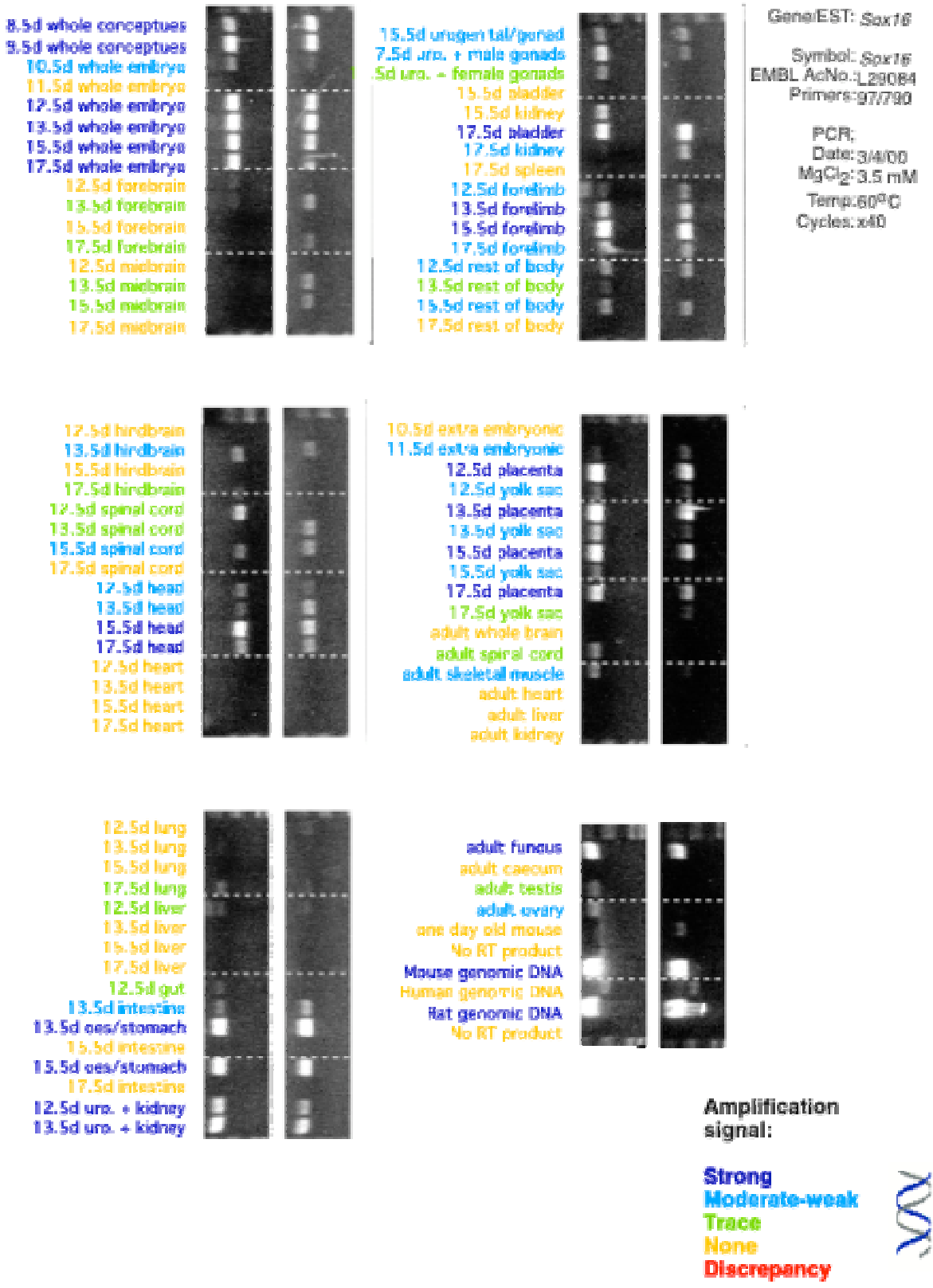
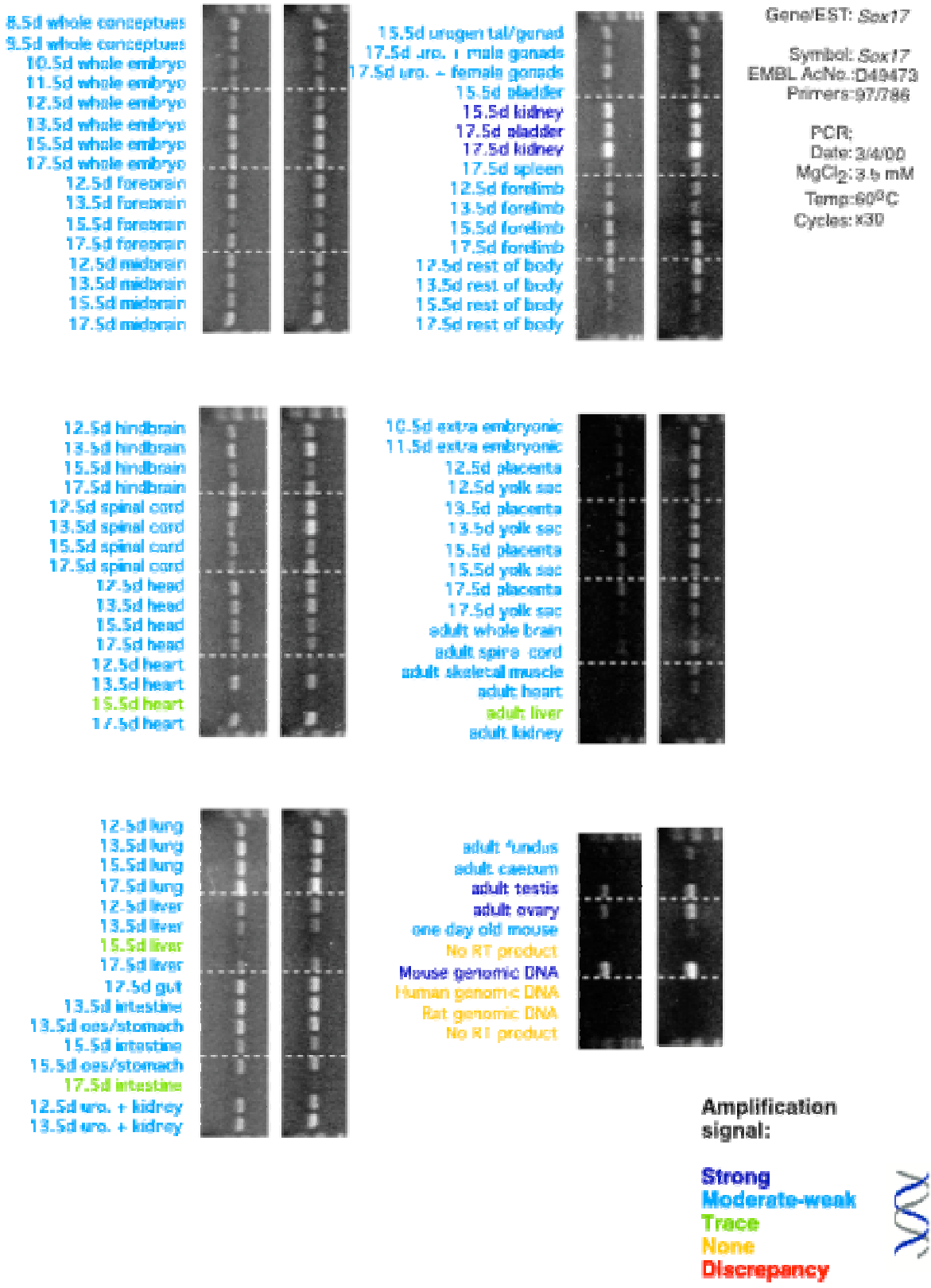


Figure 56: Expression Profile of *Sox17* in Mouse Foetal Panel



Sox17

Figure 56: *Sox 17* expression is found in all tissues, strong in adult ovary and testis, plus foetal tissues 15.5d kidney, 17.5d kidney and bladder.

The gene *Sox17* has been shown to be implicated in mouse spermatogenesis [37], which fits nicely with the strong expression found in adult testis. Another important role for *Sox17* has been identified in early mouse endoderm development [38], which might explain the general expression found through out foetal development in this study.

Sox19

Figure 57: *Sox19* in this study showed a poor expression profile and I have failed to score the results for this reason. It is included in the study to illustrate results, which require a primer redesign.

SoxLZ

Figure 58: of *SoxLZ* was found to be strongly expressed in most tissues, from foetal and adult stages. Tissues, which were particularly light included 15.5d heart, 17.5d intestine, 15.5d bladder, 17.5d spleen and the placental samples for all stages.

Slc17a2

Figure 59: *Slc17a2,* Solute carrier type 17 2a is a renal sodium phosphate/hydrogen co-transporter previously found to be highly specific for adult kidney [2], found here expressed in the 15.5d and 17.5d kidney and spleen of the 17.5d, with trace amounts in the 15.5d whole embryo and 13.5d hindbrain.

Figure 57: Expression Profile of *Sox19* in Mouse Foetal Panel

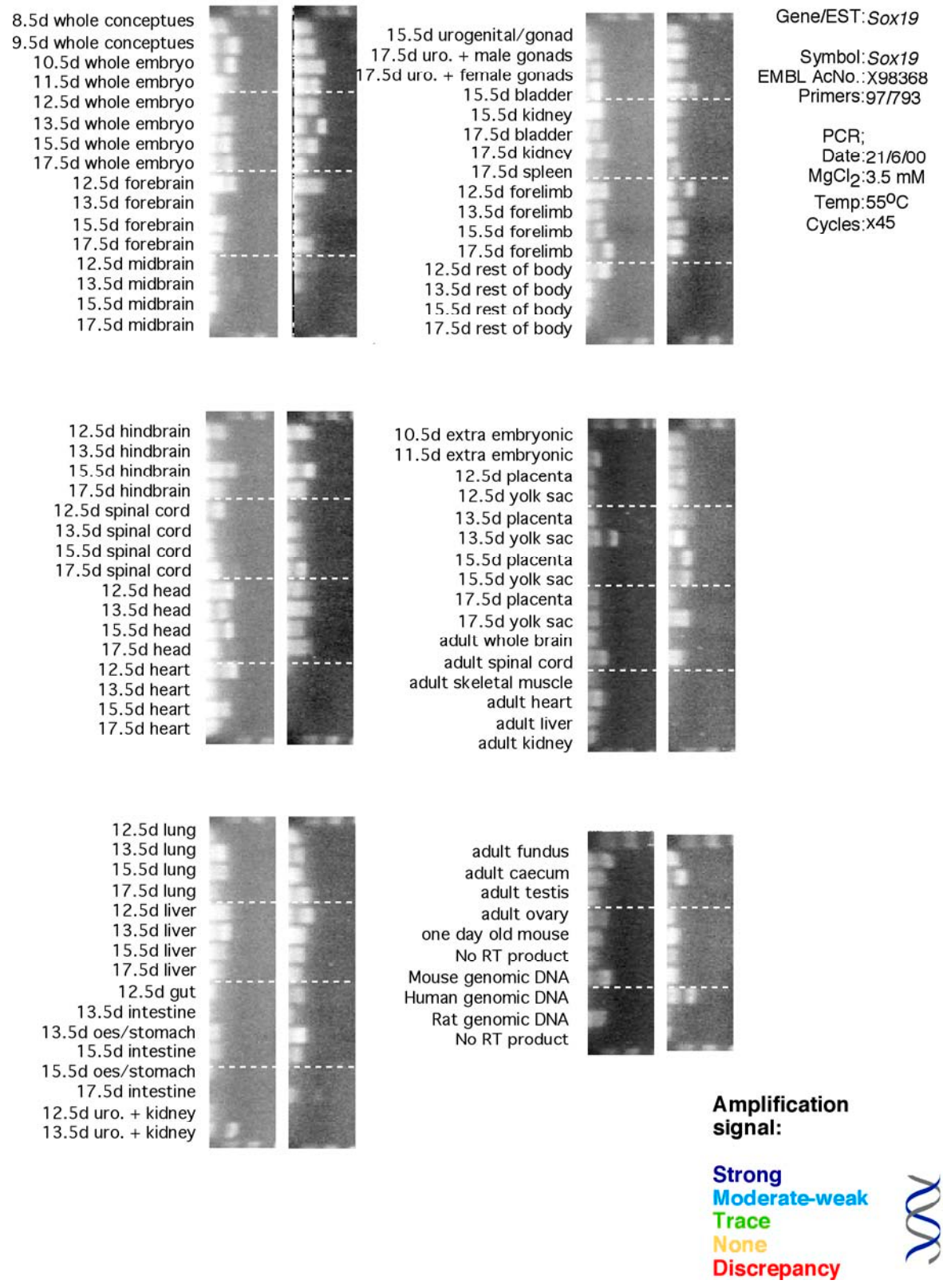


Figure 58: Expression Profile of *SoxLZ* in Mouse Foetal Panel

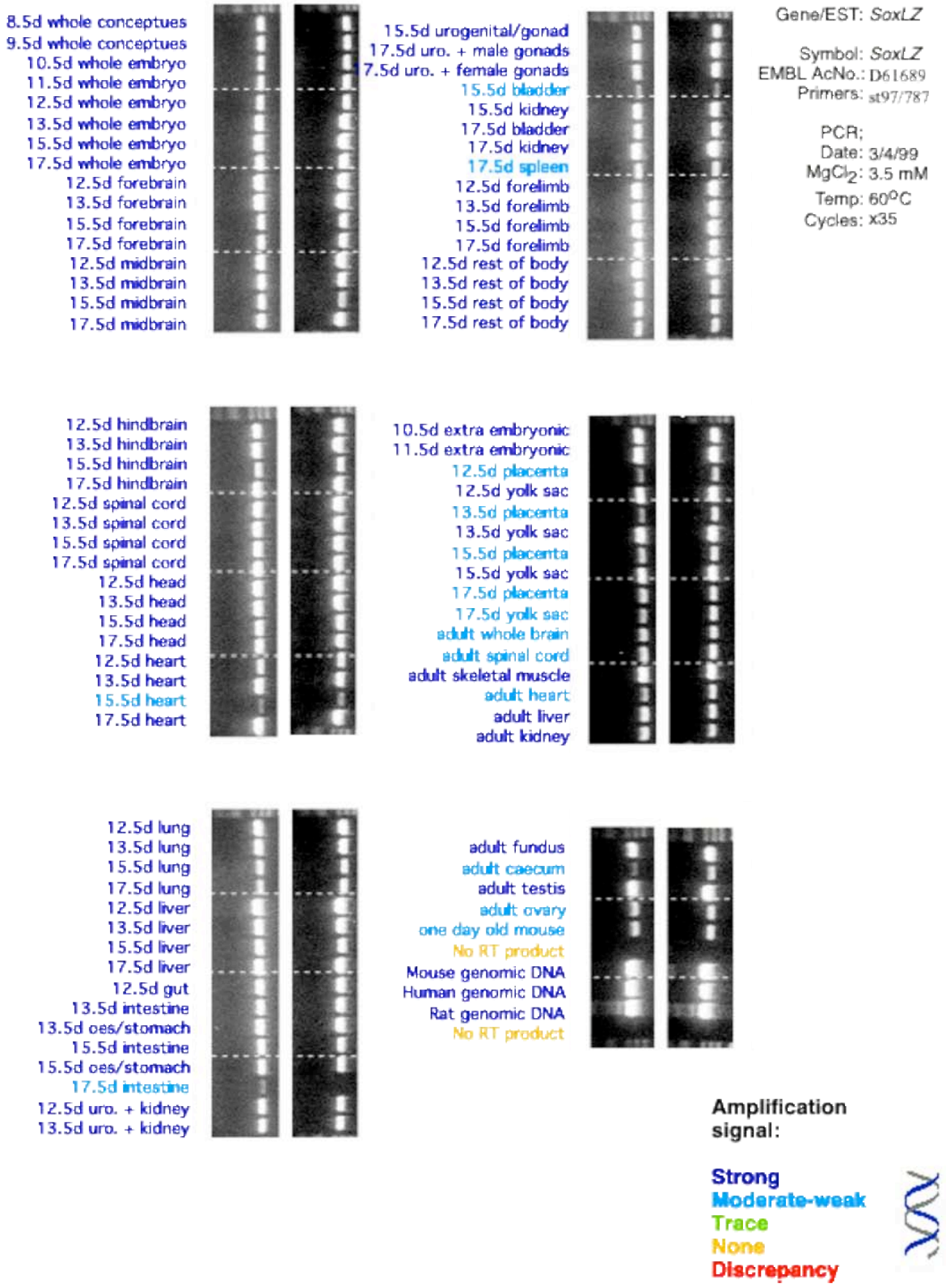
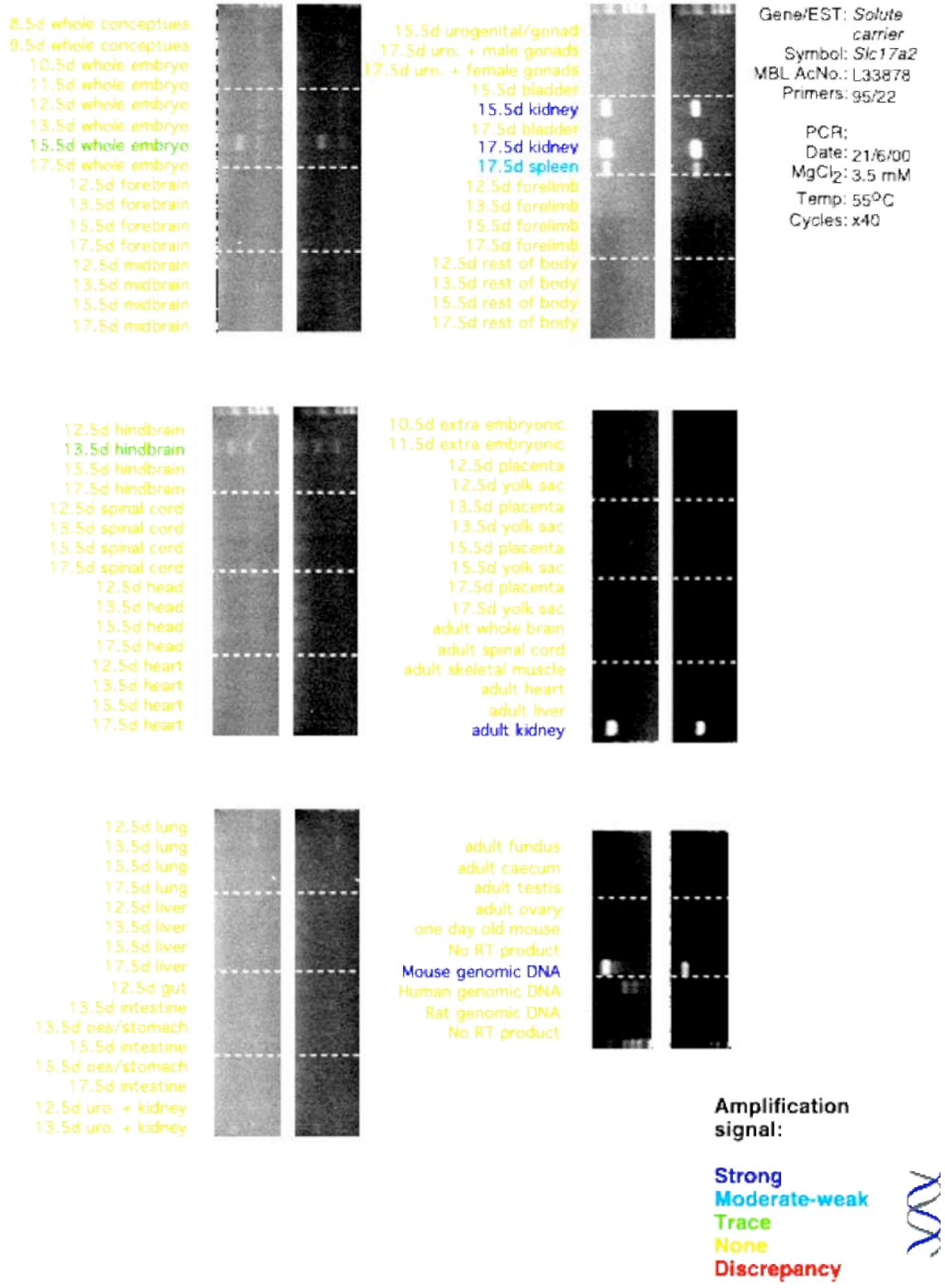


Figure 59: Expression Profile of *Solute Carrier type 17a2* in Mouse Foetal Panel



Csna1

Figure 60: *Csna1*, Alpha Casein found here only in the liver and midbrain for 17.5d had in an earlier study been found in the pregnant mammary gland and mid term foetus only [2]. This gene is associated with the metabolism of casein found in milk and so was not expected to be found in very early stages.

FABPi

Figure 61: *FABPi,* Fatty Acid Binding Protein intestinal in this study was found in trace amounts in the 12.5d and 17.5d whole embryo with higher amounts in the 15.5d whole embryo. High expression was also found in stomach/intestinal regions for all stages, with lower amounts in the extra embryonic tissues and yolk sacs of later stages. In the adult section, this gene was found at high levels in the liver and caecum confirming our earlier study [2].

Figure 60: Expression Profile of *Alpha Casein* in Mouse Foetal Panel

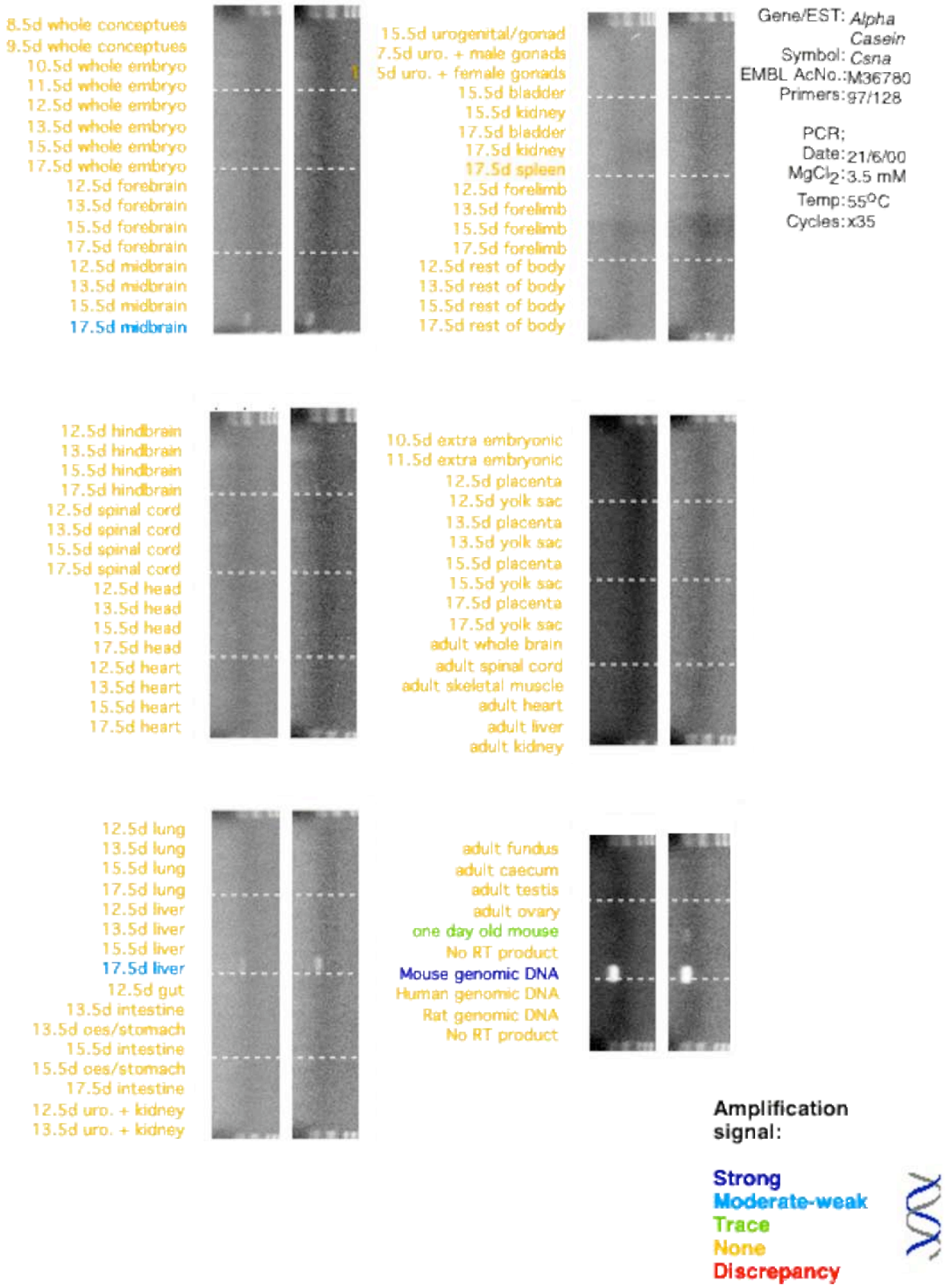
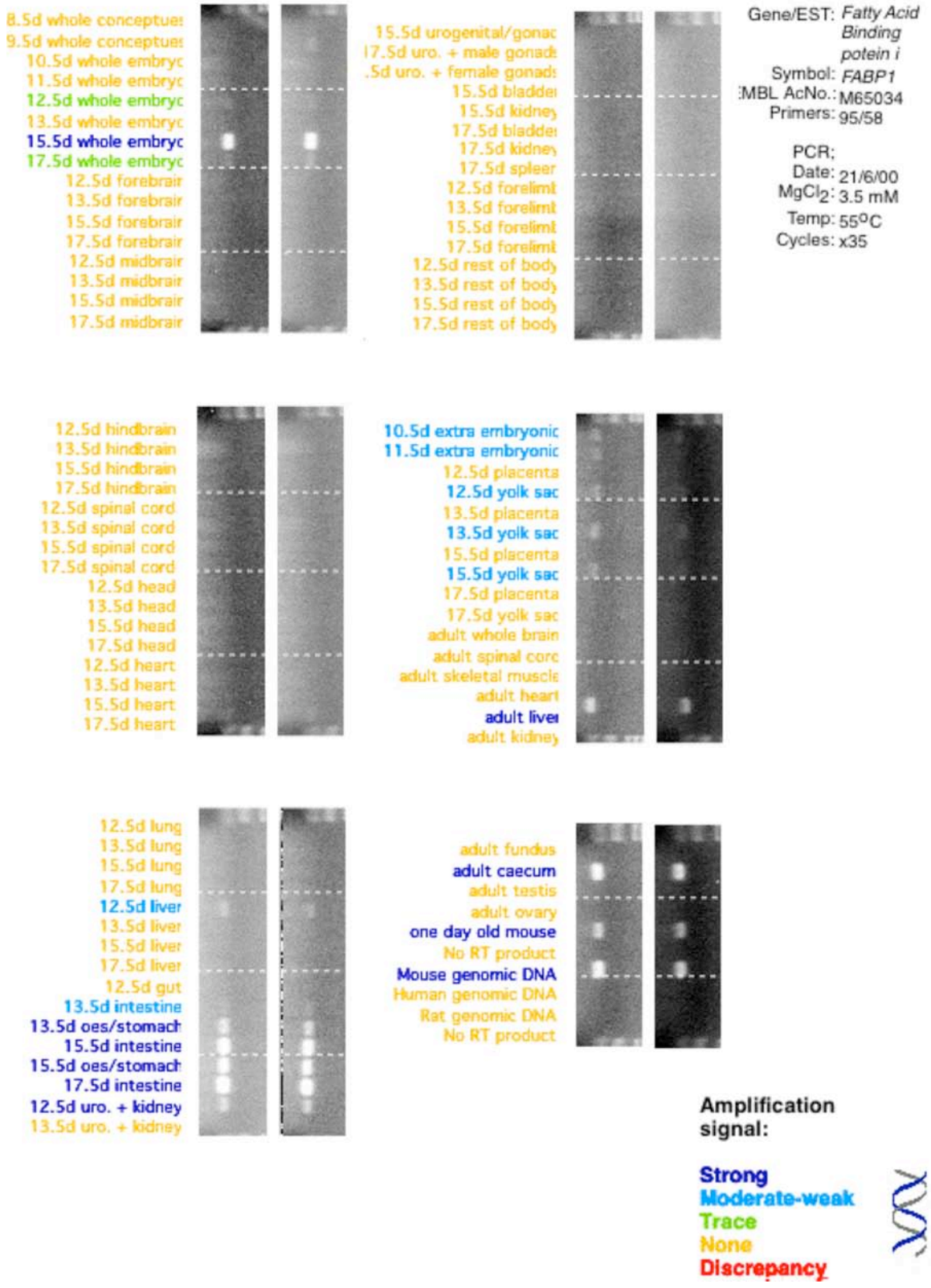


Figure 61: Expression Profile of *Fatty Acid Binding Protein intestinal* in Mouse Foetal Panel



2.4.2. Concluding Remarks

Profiles of gene expression from these tissues illustrated as agarose gels, provide some information. However, a clearer interpretation of this information is obtained, when these gels are interrogated computationally and presented as graphs, illustrating a broader range of expression intensities across the various tissues. For a complete and thorough analysis, a comparison of the individual tissue expression of the housekeeping genes as a direct relationship to each of the genes profiled would I believe provide a more accurate measure of the genes in this study. Modern methods of gene expression tend to include this form of relationship, providing a more uniform base line from which results are then presented.

The following profiles were chosen, as sufficiently different from each other, for further investigation by *in-situ* hybridization, *Sox 2, 4, 6, 15, 16, 17* and *LZ2*, in order to identify the sub-cellular localisation of the gene expression illustrated in this study.

References

1. Sonnhammer EL and D. R., *A workbench for large-scale sequence homology analysis*. Comput Appl Biosci., 1994. **10**(3): p. 301-307.
2. Freeman, T.C., et al., *Expression Mapping of Mouse Genes*. MGI Direct Data Submission, 1998.
3. Malas, S., et al., *Cloning and mapping of the human SOX1: a highly conserved gene expressed in the developing brain*. Mamm Genome, 1997. **8**(11): p. 866-8.
4. Kamachi, Y., et al., *Involvement of Sox1, 2 and 3 in the early and subsequent molecular events of lens induction*. Development, 1998. **125**(13): p. 2521-2532.
5. Pevny, L.H., et al., *A role for SOX1 in neural determination*. Development, 1998. **125**(10): p. 1967-1978.
6. Avilion, A.A., et al., *Multipotent cell lineages in early mouse development depend on SOX2 function*. Genes & Development, 2003. **17**(1): p. 126-140.
7. Ashery-Padan, R., et al., *Pax6 activity in the lens primordium is required for lens formation and for correct placement of a single retina in the eye*. Gens Dev, 2000. **14**(21): p. 2701 - 2711.
8. Kamachi, Y., et al., *Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development*. GENES & DEVELOPMENT Osaka Univ, Inst Mol & Cellular Biol, Osaka 5650871, Japan, 2001. **15**(10): p. 1272-1286.
9. Muta, M., et al., *Distinct roles of SOX2, Pax6 and Maf transcription factors in the regulation of lens-specific delta 1-crystallin enhancer*. GENES TO CELLS Japan Osaka Univ, Grad Sch Frontier Biosci, Dev Biol Lab, Osaka, Japan, 2002. **7**(8): p. 791-805.
10. Payen, E., et al., *The ovine SOX2 gene: Sequence, chromosomal localization and gonadal expression*. Gene, 1997. **189**(1): p. 143-147.
11. Soriano, N.S. and S. Russell, *The Drosophila SOX-domain protein Dichaete is required for the development of the central nervous system midline*. Development, 1998. **125**(20): p. 3989-3996.

12. Rex, M., et al., *Dynamic expression of chicken Sox2 and Sox3 genes in ectoderm induced to form neural tissue*. Developmental Dynamics, 1997. **209**(3): p. 323-332.
13. Cheung, M., et al., *Roles of Sox4 in central nervous system development*. Brain Res Mol Brain Res, 2000. **79**(1-2): p. 180-191.
14. Smits, P., et al., *The Transcription Factors L-Sox5 and Sox6 Are Essential for Cartilage Formation*. Dev. Cell, 2001. **1**: p. 277-290.
15. Chimal-Monroy, J., et al., *Analysis of the molecular cascade responsible for mesodermal limb chondrogenesis: sox genes and BMP signaling*. Dev. Biol, 2003. **257**(2): p. 292-301.
16. Blaise, R., et al., *Testis expression of hormone-sensitive lipase is conferred by a specific promoter that contains four regions binding testicular nuclear proteins*. Journal of Biological Chemistry, 1999. **274**(14): p. 9327-9334.
17. Wunderle, V.M., et al., *Cloning and characterization of SOX5, a new member of the human SOX gene family*. Genomics, 1996. **36**(2): p. 354-8.
18. Bowles, J., G. Schepers, and P. Koopman, *Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators*. Developmental Biology, 2000. **227**(2): p. 239-255.
19. Taniguchi, K., et al., *Isolation and characterization of a mouse SRY-related cDNA, mSox7*. Biochimica Et Biophysica Acta-Gene Structure and Expression, 1999. **1445**(2): p. 225-231.
20. Takash, W., et al., *SOX7 transcription factor: sequence, chromosomal localisation, expression, transactivation and interference with Wnt signalling*. Nucleic Acids Research, 2001. **29**(21): p. 4274-4283.
21. Sock, E., et al., *Idiopathic Weight Reduction in Mice Deficient in the High-Mobility-Group Transcription Factor Sox8*. Molecular and Cellular Biology, 2001. **21**(20): p. 6951-6959.
22. Schepers, G., et al., *SOX8 is expressed during testis differentiation in mice and synergises with SF1 to activate the Amh promoter in vitro*. Journal of Biological Chemistry, 2003. **278**(30): p28101-28108 **M304067200**.
23. Kent, J., et al., *A male-specific role for SOX9in vertebrate sex determination*. Development, 1996. **122**: p. 2813 - 2822.

24. Scherer, G., *The molecular genetic jigsaw puzzle of vertebrate sex determination and its missing pieces*. Novartis Found Symp 2002, 2002. **244**: p. 225- 236.
25. Bridgewater, L.C., et al., *Adjacent DNA sequences modulate Sox9 transcriptional activation at paired Sox sites in three chondrocyte-specific enhancer elements*. Nucleic Acids Research, 2003. **31**(5): p. 1541-1553.
26. Ringwald M, et al., *The mouse gene expression database*. Nucleic Acids Research, 2001. **29**: p. 98 - 101.
27. Pusch, C., et al., *The SOX10/Sox10 gene from human and mouse: sequence, expression, and transactivation by the encoded HMG domain transcription factor*. Human Genetics, 1998. **103**(2): p. 115-123.
28. Kuhlbrodt, K., et al., *Sox10, a novel transcriptional modulator in glial cells*. Journal of Neuroscience, 1998. **18**(1): p. 237-250.
29. Uwanogho, D., et al., *Embryonic expression of the chicken Sox2, Sox3 and Sox11 genes suggests an interactive role in neuronal development*. Mech Dev, 1995. **49**(1-2): p. 23-36.
30. Wright, E.M., B. Snopek, and P. Koopman, *Seven new members of the Sox gene family expressed during mouse development*. Nucleic Acids Res, 1993. **21**(3): p. 744.
31. Schepers, G.E., R.D. Teasdale, and P. Koopman, *Twenty pairs of Sox: Extent, homology, and nomenclature of the mouse and human sox transcription factor gene families*. Developmental Cell, 2002. **3**(2): p. 167-170.
32. Jay, P., et al., *SOX22 is a new member of the SOX gene family, mainly expressed in human nervous tissue*. Hum Mol Genet, 1997. **6**(7): p. 1069-77.
33. Kido, S., et al., *Cloning and characterization of mouse mSox13 cDNA*. Gene, 1998. **208**(2): p. 201-6.
34. Hargrave, M., et al., *The HMG box transcription factor gene Sox14 marks a novel subset of ventral interneurons and is regulated by Sonic hedgehog*. Developmental Biology, 2000. **219**(1): p. 142-153.
35. Beranger, F., et al., *Muscle differentiation is antagonized by SOX15, a new member of the SOX protein family*. Journal of Biological Chemistry, 2000. **275**(21): p. 16103 - 16109.

36. Hiraoka, Y., et al., *Isolation and expression of a human SRY-related cDNA, hSOX20*. Biochimica Et Biophysica Acta-Gene Structure and Expression, 1998. **1396**(2): p. 132-137.
37. Kanai, Y., et al., *Identification of two Sox17 messenger RNA isoforms, with and without the high mobility group box region, and their differential expression in mouse spermatogenesis*. Journal of Cell Biology, 1996. **133**(3): p. 667-681.
38. Kanai-Azuma, M., et al., *Depletion of definitive gut endoderm in Sox17-null mutant mice*. Development, 2002. **129**: p. 2367-2379.

CHAPTER 3

Validation of Sox Genes PCR data by *in-situ* hybridisation

Introduction

Having tested the mouse foetal cDNA panel by PCR with primers for members of the *Sox* gene family, a selection of *Sox* genes, which showed differing RTPCR expression patterns, was then taken for analysis by *in-situ* hybridization, in order to identify expression at the cellular level. Tissue sections were probed with radiolabelled oligonucleotides designed to *Sox 2, 4, 6, 15, 16, 17* and *LZ2*, together with control oligo probes for *Ribosomal protein S29* (*Rps29*), *RAR-related orphan receptor alpha* (*Rora*), *Epidermal growth factor* (*Egf*), and the sense oligo probe for *Protease 26S subunit ATPase,* (*Psmc1*).

3.1 Special Equipment and Suppliers

Sterilizing oven 180^{o}C	Scientific Laboratory Suppliers
Incubating oven 40-60^{o}C	Scientific Laboratory Suppliers
Cryostat	Leica
Autoradiography equipment and dark room facilities	
Safe light filter	Kodak 1 # 152 1509, Ilford 914 # CEA 4B, or Agfa-Gaevert # R1
X-ray film	Kodak BioMax MR # 871 5187, # XAR 5 or Amersham Hyperfilm β-max
Rapid fix	Kodak

X-ray cassette	Amersham Pharmacia
X-ray film developer	Kodak # D19 or Ilford#D19
X-ray envelope	Amersham Pharmacia
Liquid photographic emulsion	Kodak # NTB-2 or Ilford#K5
LC29	Ilfotec # LC29
Superfrost slides	BDH # 406/0169/02
Scotlab Easicounter 4000	
Hot dot: HCT.dot, sample range 0.5 – 2 □l	SL 8329 # B-QC-7001, 48
Nunc square petri dish	Nunc # 166508
Slide racks stainless steel or black plastic-style	BDH
Cryo-M bed	Bright Instruments
Nuclear research emulsion in gel form	Ilford
Variable light intensity illuminator, dual lamp precision	Northern Light
Microscope Leica DM RB	Leica
Lenses x5, X10, X20, X40	Leica
Lens, Nikon PN-11 52.5	Nikon
Lens, NIKKOR 105mm 1:2,8	Nikon #244228
Relay camera MTI CCD 72	MTI
Leica DC 200 version 2.51 windows NT application	Leica
MCID Image Analyser software, M4-Image version 3, Rev 1.4	Imaging Research Inc. Brock University, St. Catherines, Ontario, Canada L2S 3A1

3.2. Materials and Solutions

Chemicals and Suppliers

"(poly)L-lysine hydrobromide (mol. wt > 350,000) "	Sigma # P1524
[35S]deoxyadenosine ‘5 (a-thio)triphosphate 10 mCi/ml	Amersham # SJ1334
amberlite mixed bed resin MB-150	Sigma # A5710
ammonium acetate	BDH # 100134T
aluminium sulphate	Sigma # A 7523
bovine serum albumin fraction V	Sigma # A8022
C14 standards	Amersham # RPA 504
chromium potassium sulphate	Sigma # C 5926
dextran sulphate	Pharmacia # 17-0340-02
dithiothreitol	Sigma # D9779
DPX mountant	BDH # 360294H
Ethanol	BDH # L227007
ethylene diamine tetraacetic acid (EDTA)	Sigma # ED2SS
Eosin yellowish	BDH # 341972Q
ethylene glycol	Sigma # E-9129
ficoll 400	Sigma # F2637
formamide super pure	Gibco/BRL # 540-5515
glass wool	Sigma # G8389
glycerol	Sigma # G 7757
hematoxylin	Sigma # H 9627
heparin (sodium)	BDH/Merck # 28470
isopentane	Sigma # 27,041-5
paraformaldehyde	Merck # 21046205 438

phosphate buffered saline (PBS)	Oxoid #BR14
polyadenylic acid (potassium salts)	Sigma # P9403
polyvinylpyrrolidone	Sigma # P5288
sephadex G50 medium	Pharmacia # 17-00 45-01
"sheared salmon sperm DNA 10 mg/ml type III, sodium salt, "	Sigma # D-9156
sigmacote siliconising solution	Sigma # SL2
sodium azide 1% w/v	Severn Biotech
sodium chloride	BDH # 102415K
sodium hydroxide	BDH # 102525P
sodium dodecyl sulphate (SDS)	BDH # 442444H
sodium dihydrogen orthophosphate	BDH # 301324G
disodium hydrogen orthophosphate	BDH # 102494C
sodium iodate	Sigma # S 4007
sodium metabisulphite	Sigma # S 1516
sodium pyrophosphate	Sigma # S 6422
sodium thiosulphate 5 hydrate	BDH # 102684G
terminal deoxynucleotidyl transferase	Pharmacia # 27-0730-0
trisodium citrate	BDH # 102425M
xylene	BDH # 102936H

3.2.1. 10 X Phosphate-buffered saline (PBS) *fc*(final concentration)

75.972 g sodium chloride *1.3 M*

9.937g disodium hydrogen orthophosphate *70 mM*

4.14g sodium hydrogen orthophosphate *30 mM*

made up to 1 L with sterile water, treated with 1 ml depc and autoclaved.

1 X PBS was prepared by diluting the stock in depc-treated water.

3.2.2. (Poly)L-lysine *fc*

25 mg (poly)L-lysine hydrobromide (mol. wt > 350,000) *5 mg/ml*

5 ml of depc-treated water

Store frozen as 1 ml aliquots in sterile eppendorf tubes

3.2.3. Ethanol

Absolute (99.7%) ethanol is diluted with depc-treated water to give the required concentration (for example, 50%, 70%, 95%).

3.2.4. 4% Paraformaldehyde in PBS

Paraformaldehde powder is toxic and harmful to mucous membranes: thus always wear gloves and a facemask when handling it and perform the following procedure in a fume hood.

- 40 g paraformaldehyde, slowly added to 500 ml depc-treated water in a sterile 1 L beaker.
- Heated to 60^{o}C with continuous stirring, using an autoclaved magnetic stir bar: heat kept below 65^{o}C.
- Drops of 6 M NaOH gradually added to the milky suspension until it clears.

- 100 ml sterile 10 x PBS in depc-treated water, then added and mixed thoroughly.
- Cooled and made up to volume in a 1 L volumetric flask.
- PH checked as 7.0 with pH paper, and stored in a refrigerator/cold room at 4^{o}C.
- Paraformaldehyde solution is best used within a 2 week period, since it polymerises.

3.2.5. 20 x SSC *fc*

438.25 g sodium chloride *3.75 M*

220.5 g trisodium citrate *0.375 M*

2 L water

2 ml depc *0.1%*

treat as depc water, see section 1.2.1.

3.2.6. Dithiothreitol

Dithiothreitol (DTT) is a reducing agent and when using ^{35}S-labelled probes, it will reduce or prevent cross-linking of sulphur residues *fc*

3.09 g DTT *1 M*

29 ml depc-treated water.

Stored as 1 ml aliquots in sterile eppendorf tubes at -20^{o}C.

3.2.7. TENS buffer

To a sterile 500 ml bottle the following are added: *fc*

20 ml 1 M Tris-HCl (pH 7.5) *20 mM*

5 ml 0.5 M EDTA (pH 7) *5 mM*

14 ml 5 M sodium chloride *140 mM*

5 ml 10% sodium dodecyl sulphate *0.1%*

466 ml depc-treated water

Filter through a 0.2 μm filter into a sterile baked bottle.

3.2.8. Sephadex G50 solution

In a sterile 250 ml glass bottle the following are added: *fc*

5 g Sephadex G50 medium *24%*

120 ml filtered TENS buffer

- left to swell for 1-2 hours.
- Washed several times with TENS buffer to remove soluble dextran and 'fines'.
- Autoclaved at 10 lb/sq. inch (note the lower pressure) for 20 mins.
- TENS supernatant removed and sephadex store at + 4^{o}C with 5 mg of sodium azide (this should last a year, however, always check for evidence of microbial growth).

3.2.9. Tailing buffer

10 x Tailing buffer is obtained from Pharmacia. It is supplied free of charge with purchase of terminal deoxynucleotidyl transferase enzyme. 10 x Tailing buffer is referenced Deng, G, R and Wu, R., Methods in Enzymology, 100, 96 (1983).

3.2.10. Heparin *fc*

120 mg heparin (sodium) *120 mg/ml*

1 ml depc-treated water.

Store frozen in 50 μl aliquots in sterile eppendorf tubes.

3.2.11. Polyadenylic acid *fc*

100 mg polyadenylic acid (potassium salts) *10 mg/ml*

10 ml depc-treated water.

Stored frozen in 1 ml aliquots in sterile eppendorf tubes.

3.2.12. 50 X Denhardt's Solution *fc*

5 g Ficoll 400 *1%*

5 g polyvinylpyrrolidone *1%*

5 g bovine serum albumin fraction V *1%*

500 ml depc-treated water

Stir until all are dissolved completely. *Stored frozen as 25 ml aliquots.*

3.2.13. 0.5 M diSodium Hydrogen Orthophosphate

3.559 g disodium hydrogen orthophospate

50 ml depc-treated water

3.2.14. 0.5 M Sodium diHydrogen Orthophosphate

3.45 g sodium dihydrogen orthophosphate

50 ml depc-treated water

3.2.15. 0.5 M Sodium Phosphate pH 7 *fc*

7.8 ml depc-treated 0.5 M disodium hydrogen orthophosphate *0.5 M*

12.2 ml depc-treated 0.5 M sodium dihydrogen orthophosphate *0.5 M*

pH checked as 7

3.2.16. 0.1 M Sodium Pyrophosphate

2.23 g sodium pyrophosphate

50 ml depc-treated water

3.2.17. Oligonucleotide Hybridization Buffer

		fc
25 ml	thawed deionized formamide	*50%*
5 g	Dextran Sulphate	*10%*
10 ml	20 x SSC	*4 x SSC*
5 ml	Denhardt's solution	*5 x*
10 mg	Salmon sperm DNA	*0.02%*
1 ml	5 mg/ml polyadenylic acid	*0.01%*
50 µl	120 mg/ml heparin	*0.012%*
2.5 ml	0.5 M Sodium Phosphate pH 7	*0.025 M*
0.5 ml	0.1 M Sodium Pyrophosphate	*0.01 M*

- Left overnight on a rotary mixer, at room temp.
- Made up to 50 ml with depc-treated water, wrapped in foil and stored at 4^{o}C.
- On use, left at room temp for 30-60 mins and inverted several times, to allow bubbles to rise. Hybridization buffer should be okay for up to a year.

3.2.18. Deionized Formamide

Formamide is a potential carcinogen, wear gloves and do all procedures in the fume hood.

2 x 100 ml thawed super pure formamide (stored -20^{o}C)

200 g mixed bed resin Amberlite MB-150

stirred for 30mins

pH checked as approx 7

filtered and dispensed as 25 ml aliquots in 50 ml falcon tubes, store at -20^{o}C.

3.2.19. *In-situ* wetting solution	*fc*
50 ml 20 x SSC	*4 x SSC*
125 ml formamide	50%
75 ml double deionied water	

3.2.20. Wash solution	*fc*
50 ml 20 X SSC	*1 x SSC*
2 g sodium thiosulphate 5 hydrate	0.2%
1 L double deionised water	

3.2.21. Ammonium acetate solution	*fc*
11.56g ammonium acetate	*600 mM*
50 ml glycerol	*0.5%*
950 ml double deionised water	

3.2.22. Developer

10 ml Ilford LC29

290 ml double distilled water

3.2.23. Stop solution *fc*

6 g sodium metabisulphate *2%*

6 g chromium (III) potassium sulphate *2%*

300 ml double distilled water

3.2.24. Fix solution

100 ml B&W fixer

200 ml double distilled water

3.2.25. Hematoxylin *fc*

6 g haematoxylin *0.6%*

4.2 g aluminium sulphate *0.42%*

1.4 g trisodium citrate *0.14%*

0.6 g sodium iodate *0.06%*

269 ml ethylene glycol 35%

680 ml double distilled water

3.2.26. 1% acid ethanol

1% conc. Hydrochloric acid in 100% ethanol

3.2.27. 1% Eosin

1 g Eosin Y

100 ml double distilled water

3.2.28. Siliconizes glass wool

- Place glass wool in an autoclavable bottle.
- Pour in sigmacote
- Invert bottle a few times and remove the sigmacote.
- Autoclave.

3.3. Methods

General notes: RNA, unlike DNA, is highly susceptible to degradation, the most common cause being RNases found in abundance in bodily secretions (sweat and saliva). RNA is also susceptible to heat denaturation. Therefore it is strongly recommended that all glassware is baked, plastics autoclaved, solutions treated with depc where possible and gloves worn throughout.

All solutions must be free of RNase and stored in sterile (RNase-free) containers. Gloves must be worn at all times when handling containers. Also, all solutions must be treated with depc (0.1%) and then autoclaved: stock solutions are diluted with depc-treated water. Diethylpyrocarbonate (depc) is an alkylating agent and inactivates any protein present in the solutions.

3.3.1. Preparation of Glass Slides

Proper pre-treatment of glass slides is necessary to avoid:

- *sticking of labelled probe to glass and causing undesired background,*
- *degradation of mRNA,*
- *sections falling off during the hybridisation and post-hybridisation procedures.*
  - Good quality, precleaned slides, such as BDH Superfrost. 0.1mm thick were used. Wearing gloves throughout, the cellophane was removed and wrapping them in batches of 50, in aluminium foil, and baked in an oven at 180^{o}C for 4 hours, on a shelf, to destroy the ribonucleases.
  - Slides were removed from the oven and allowed to reach room temperature slowly.
  - 1 ml (poly)L-lysine hydrobromide (5mg/ml) was allowed to thaw and diluted in 50 ml with depc-treated water (to obtain 0.01% (poly)L-lysine) in a sterile 50

ml Falcon propylene; mixed well and poured into a square sterile Sterilin petri dish.

- Completely immersing each glass slide (one at a time) into the (poly)L-lysine solution for approximately 5 seconds; removed with another glass slide, taking care to only handle from the edge, the excess (poly)L-lysine solution was drained by blotting the bottom of the slide with a paper towel, placing in a sterile slide rack, and allowed to air-dry in a dust-free area, covered loosely with foil.
- When the slides were dry, they were placed in a slide box with silica gel and stored at 4^{o}C until used. Usable up to 1 month after coating with (poly)L-lysine, but is advisable to coat slides immediately prior to use, and not store for prolonged periods.
- 300 are an optimal number to process.

3.3.2. Collection of Tissues

Procedures must be adopted to minimize the effects of stress on the animal both in terms of animal welfare and the possible changes in gene expression, in addition to the possible degradation of mRNA. Therefore, animals must be killed quickly and the tissues removed and frozen as rapidly as possible without damage.

Tissues of interest were removed directly into isopentane (5 secs) held on cardice. Small samples were embedded directly in mounting medium e.g. DPX mountant or Cryo-M Bed, Bright Instruments. Samples were then stored in liquid nitrogen or at –70^{o}C.

3.3.2.1. Sectioning of Tissues

- Prior to cutting sections, tissues were transferred to the cryostat and allowed to equilibrate to the temperature of the cryostat.
- The cryostat microtome knife was sterilised with a piece of tissue soaked in absolute alcohol.
- Tissue was mounted on to the block using mounting medium, and the block held in the chuck ensuring the block was orientated to give the desired plane of sectioning.
- The block was trimmed until the area of interest was reached.
- The temperature of the chuck and the chamber may be altered depending on the tissue, and this is largely trial and error. In general, however, a temperature of between -15^{o}C and -20^{o}C is fine for most tissues. Sections may require cutting slowly or quickly, depending on tissue and temperatures. After cutting, minimal manipulation was used (merely teasing out with a fine paint brush), prior to placing a slide close enough for the section to 'jump' onto the coated slide in the orientation required to display a fully stretched section.
- Sections were cut at 10 μm thickness precisely. Multiple sections of various stages to a slide. Developmental stages 8.5d, 10.5d, 12.5d, 13.5d, 15.5d, and 17.5d were cut over a period of a few weeks, processing a day's worth of cutting at any one time.
- Once the slide had a full complement of sections on, it was placed on a rack, face away to thoroughly air dry for 1–2 hrs prior to fixing.
- Sections were fixed on the day of sectioning and this step required at least 1 hr. All sections were fixed on the day of cutting and stored at –70^{o}C until used.

- If it is impossible to fix on the day of sectioning, it is acceptable to place the slides in a slide box with silica gel, seal with tape and leave at –70^{o}C overnight. Next day remove the slide box from the -70^{o}C, allow to reach room temperature, then fix the sections in 4% paraformaldehyde.

3.3.2.2. Fixation of sections

Efficient fixation procedures are important for maintaining good tissue morphology, for stabilization and retention of tissue mRNAs, and for destroying residual RNase. Over-fixation may inhibit accessibility of probe to mRNA. Several pre- and post-fixation procedures are available.

The slides were placed (with mounted sections) in sterilized RNase-free autoclaved stainless steel or black plastic-style racks and treated as follows:

- 5 mins in 4% paraformaldehyde in 1 x PBS, chilled on ice.
- 2.5 mins in 1 X PBS.
- 2.5 mins in 1 X PBS.
- 5 mins in 70 % ethanol/depc-treated water.
- 5 mins in 95 % ethanol/depc-treated water.

3.3.2.3. Storage of fixed sections

After fixation, the sections were stored in clean slide boxes either at -70^{o}C or in 95 % ethanol/ depc-treated water at 4^{o}C in a spark-proof cold room or refrigerator. Sections held at -70^{o}C can lead to tissue desiccation and thus poor histology and mRNA signal, moisture condensation may occur on the sections on removal from -70^{o}C if sections are not allowed to come to room temperature properly. Moisture may

cause leakage of cellular RNase and thus mRNA degradation. These problems are not encountered when sections are stored in alcohol at 4^{o}C. Storage in alcohol in addition to the easy access and visualization of the appropriate sections needed for *in situ* hybridization also helps to de-fat the sections. This latter point is important since some probes bind to the white matter and may give increased background/non-specific labelling. Under these conditions, mRNA is stable for years in 95% ethanol/depc-treated water at 4^{o}C.

3.3.3. Design, synthesis and purification of oligonucleotide probes

Various types of nucleic acid probes have been used to detect mRNAs in tissues by *in-situ* hybridization, for example, ss (single stranded) c (complementary) DNA, ds cDNA, cRNA and synthetic oligodeoxyribonucleotide probes. The advantage of using oligoprobes are that they can be synthesised in large quantities to an accurately defined sequence, are more stable than other probes, can differentiate single base differences and being small molecules, have good tissue penetration properties. However, they are less sensitive than RNA probes and the hybridisation conditions are critical to a successful result. As the expertise to work with oligo probes existed in the group, it was decided to follow this route. Several factors must be considered when designing oligonucleotide probes. Probe length and GC/AT ratio are the most important aspects since they may affect the stability of the hybrid. In general, probes with greater GC content will form more stable hybrids, since GC base pairs are stabilized by three hydrogen bonds in contrast to the two hydrogen bonds that stabilize AT/UT base pairs. However, if the GC content of the oligonucleotide is too high (>65%), non-specific labelling may occur since the thermal stability of the probe itself will be greater. The probe length and GC content, in addition to the formamide

and salt concentrations in the hybridization buffer, will also determine the appropriate conditions for hybridization and post-hybridization treatments.

Good results were obtained using the Oligo Primer Analysis Software version 5.1 from NBI/Genovus (http://www.natbio.com), with the following settings:

- Oligonucleotide length approximately 45 bp.
- Percentage GC content between 55-65.
- Duplex (interprobe complementarity) formation.
- Hairpin (intraprobe complementarity) formation, energies in kcal/mol need to be as near to zero as possible.
- Checked for related sequence homology by BLAST to minimize non-specific hybridization.
- Correct orientation checked prior to ordering.
- Ensuring the probe was synthesized by the highest purification procedure.
- Store stocks kept at 1 □g/ul in depc-treated water, working solutions at 5 ng/ul.

Some of the *in-situ* probes, were designed by Sarah Hunt at the Sanger Centre using the computer programme for choosing PCR and DNA sequencing primers (Hillier and Green), original default parameters assumed.

Figure 62 lists the *In-situ* Probe Sequences chosen from the PCR expression profile patterns for this study.

Figure 62: *In-situ* Probe Sequences

Gene Name	Probe sequence	Symbol	Accession number
Epidermal growth factor	GCAGGTGACTGATTTCTCCCTGAGACAGGCACAACCAGGCAAAGG	*Egf*	J00380
Protease 26S subunit, ATPase	GGCGGCACTGGGTGTGAGGTGTTACCAGTGGCAGTTTGCTGGCAG	*Psmc1*	U39302
Ribosomal protein S29	ATGTTCAGCCCGTATTTGCGGATCAGACCGTGGCGGTTGGAGCAG	*Rps29*	L31609
RAR-related orphan receptor alpha	CTTCTTCGTGACTGAGATACCTCGGCTGGAGCTCGCATAGCTCTG	*Rora*	U53228
Sox2	GCGGAGCTCGAGACGGGCGAAGTGCAATTGGGATGAAAAAACAGG	*Sox2*	U31967
Sox4	CCTAAGTCCCTTTCCTGCAGTGCAAAGCCAAAGCGACTCTGGCTC	*Sox4*	X70298
Sox6	CTCCATCCGCTGTACAGGCAAATGGAGAGGTGGCTTGCTTGGAAG	*Sox6*	U32614
Sox15	TTGGGGGGCTGGTACCCAAAGCCTCTGCTCCCTTGGGTAGTTGTG	*Sox15*	X98369
Sox16	TCCTCCACGAAGGGTCTCTTCTCTTCATCGTCCAGCAGCTTCCAC	*Sox16*	L29084
Sox17	GACCTGAGGCTCGAAAGGCTGGGGCAAGGAAGCGTCTAATGTAAC	*Sox17*	D49473
SoxLZ2	GTGAGGTTTGTTGTGCATTATGGGGTGCAGAGGCAGATGGGAGGC	*SoxLZ*	D61689

3.3.4. Radioactive labelling and purification of oligonucleotide probes.

The commonly used radioisotopes for labelling oligonucleotides are ^{32}P and ^{35}S. These produce labelled probes of high specific activity allowing for fast regional localization of mRNA in tissue sections. The ^{35}S provides probes of high specific activities, with good resolution on X-ray film and with excellent cellular resolution after liquid emulsion autoradiography. It also has a longer half-life (87.4 days) and is less hazardous to use than ^{32}P.

Tailing with terminal deoxynucleotide transferase enzyme catalyses the repetitive transfer of mononucleotide units from a deoxynucleoside triphosphate to the 3'-OH terminus of the synthetic oligonucleotide with the release of inorganic pyrophosphate. The length of the poly (dATP) tail is in the order of 15 - 25 residues and can be checked by PAGE. The number of residues transferred is influenced by enzyme concentration, oligonucleotide/isotope molar ratio, and the duration of the incubation.

3.3.4.1. Radioactive ^{35}S-labelling of oligonucleotide probes.

For each of the eleven oligonucleotide probes (see above Figure 62):

In a sterile 1.5 ml eppendorf tubes on ice the following was prepared:

- 1.25 □l 10 x One-Phor-All Buffer
- 7.25 □l depc-treated water.
- 2.0 □l (10 ng) oligonucleotide
- 1 □l terminal deoxynucleotidyl transferase enzyme
- 1 □l [^{35}S]deoxyadenosine '5 (□-thio)triphosphate

Using label as fresh as possible, ^{35}S label stored at -70 ^{o}C and thawed immediately before use, behind a screen, taking extreme care when removing the foil cover if new, the remainder subaliquoted and stored at -70^{o}C.

- Total volume of the contents of the reaction tube was checked as 12.5 µl, mix gently by pipetting up and down with a gilson, to avoid air bubbles.
- Incubated immediately in a waterbath at 32 ^{o}C for 1 - 1.5 hrs.
- The reaction was stopped with the addition of 40 µl depc-treated water.
- Labelled probe was purified using a sephadex G50 spin column. This separates unincorporated nucleotides from the labelled probe.

3.3.4.2. Purification of Probe

- The 'wings' of a disposable 1 ml syringe were clipped, and place into a sterile 15 ml falcon tube such that it did not fall into the tube and yet the lid can be satisfactorily applied.
- The syringe plunger was removed and used to plug the bottom of the syringe with autoclaved siliconized glass wool.
- Using a 5 ml Gilson, pre-swollen G50 Sephadex in TENS buffer, was quickly and gently poured into the 1 ml syringe at an angle, any bubbles were then removed by sharply flicking/shaking the syringe.
- The TENS buffer elutes into the Falcon tube as the syringe was filled with with the Sephadex G50 slurry. Removing the air bubbles by gently tapping the syringe.
- Excess TENS buffer drained into the Falcon tube and discarded.
- The whole assembly (Falcon tube and syringe) was covered with the Falcon tube cap and spun at 2000 rpm for precisely 2 mins in a low speed centrifuge.

Discarding the eluate. The packed G50 column reaching around the 0.9-1 ml mark on the syringe.

- The sephadex-filled 1ml syringe, was removed, to allow the sterile decapped eppendorf tube (appropriately labelled according to the probe used) to be placed in the Falcon tube and the Sephadex G50 syringe column replaced on top.
- 2 µl 1M DTT was added to the sterile decapped eppendorf tube, at the base of the Falcon tube.

3.3.4.3. Removal of unincorporated label

- 52.5 µl of probe solution was then pippeted onto the centre of the top of the Sephadex column and spun precisely as before (2000 rpm for 2 mins). The purified labelled probe in about 50 µl volume was collected in the decapped eppendorf tube, containing the DTT.
- Mixed by vortexing and kept on ice.
- 2 µl of probe solution was transferred to a hot dot (sample range 0.5 - 2µl, SL 8329 # B-QC-7001, 48 cont. HCT.dot.) for counting in the Scotlab Easicounter 4000. Counts per minute recorded below: usually in the range of 5,000-20,000 cpm; calculate dpm [cpm x efficiency of the counter, the Scotlab Easicounter 4000 is assumed to have an efficiency rate of 7%; thus cpm / 7 x 100 = dpm] and diluted in hybridization buffer such that 100 µl will give in the region of 80 – 400,000 counts and preferably in excess of 100,000 dpm/slide.
- Figure 64 is a table showing ^{35}S incorporation into probes, with the relevant volumes to give 100,000-200,000 dpm/slide.

Figure 63: Table of label incorporated into Probes

Probe	cpm	dpm $x10^3$	vol for 2 slides	Dpm x 103 per slide
Egf	2016	14	28□l	196
Psmc1	12435	88	4□l	176
Rps29	9201	65	6□l	195
Rora	2679	19	21□l	199.5
Sox2	6162	44	9□l	198
Sox4	704	-		
Sox6	9183	61	6□l	183
Sox15	3482	25	16□l	200
Sox16	4223	29	14□l	203
Sox17	5956	42	9□l	189
SoxLZ2	3849	27	14□l	189

- Quality of the labelling was checked prior to removing the sections from storage in alcohol and proceeding with the hybridization step.

3.3.4.4. Incubation of sections with labelled probe

- Slides with the appropriate sections were removed from the alcohol, placed in a slide rack and air-dry thoroughly.
- Labelled probe diluted in hybridization buffer to give 100 – 200,000 dpm/100 □l hybridization buffer/slide – see above.
- Slides placed in a sterile Nunc petri dish, (holds approximately 20 slides).

- 100 µl aliquots of hybridization buffer applied to each slide, as 4-5 droplets evenly across the section ensuring uniformly distributed over the section.
- Strips of parafilm were used to cover the sections (the inner surface of the Parafilm is next to the sections); avoiding air bubbles. Dispersing 4-5 droplets of the hybridization buffer over the section and stand for a few seconds before applying the Parafilm.
- The dish was humidified with a piece of Kleenex tissue soaked in 50% formamide/4 x SSC (in-situ wetting solution).
- The petri dish was wrapped with saran wrap.
- Incubated at 42^{o}C overnight.

3.3.5. Post-hybridization treatments

Labelled sections were washed to remove non-hybridized probe under stringency conditions that will reduce background without losing the signal to verify the specificity of the probe for the specific mRNA. RNase contamination is no longer a concern and aseptic techniques were not observed after this point.

3.3.5.1. Slide washing

- 2 L wash solution, placed in three sandwich boxes, maintaining one at 55^{o}C.
- Slides were lifted up from the bottom of the petri dish, using a razor blade, and one at a time, in order, placed under 1 x SSC at room temperature to remove the parafilm coverslip. The parafilm was gently lifted off, without smearing the section, and the slide washed by agitation to remove excess hybridization buffer and unhybridized probe. Slides were then placed in order, in the rack and kept under the 1 x SSC, until the full complement of slides was collected.

- The rack of slides was transferred to a box of wash solution at 55^{o}C for 30 mins.
- The rack of slides was then transferred to a fresh box of wash solution at 55^{o}C for 30 mins.
- The rack of slides was then transferred at room temperature sequentially through the following boxes:
    - 1 X SSC for 5 seconds.
    - 0.1 X SSC for 5 seconds.
    - 18 ohm water for 5 seconds.
    - 70% Ethanol for 2 seconds.
    - 95% Ethanol for 2 seconds.
- The sections were then thoroughly air-dried at room temperature before exposure to dry X-ray film.

3.3.6. Autoradiography

All autoradiographic procedures must be carried out in a light-tight dark room with suitable safelight illumination with appropriate filters and a 15 W bulb (see section 3.1)

3.3.6.1. Exposure to X-ray film

- Card was placed in a cassette, without an intensifying screen and, using double sided sticky tape, rows were created to place the dried slides, in order, side by side, in a perfectly straight line. C^{14} standards at one edge.
- In the dark room, a sheet of X-ray film was placed over the arranged slides, the cassette close securely and seams of the cassette taped with masking tape.

- The cassette was placed in a cool dark place (cupboard or drawer) away from any other source of radiation.
- The length of exposure of the labelled sections to film depends upon several factors including the specific activity of the labelled probe, and the mRNA and its abundance in the relevant sections under investigation, and varies from 3 days to 3 weeks.

3.3.6.2. Development of the X-ray film

Developed in an automatic developer at appropriate settings or as follows:

- Kodak D19 or Ilford D19 developer at 21-23^{o}C for 5 mins.
- Running water for 30 secs.
- Rapid fix (Kodak) for 5 mins.
- Wash in running cool water for 30 mins.
- Air dry for 30 - 60 mins.

Label the film carefully and store in an X-ray envelope.

3.3.7. Emulsion autoradiography

For the cellular localization of mRNA transcripts, dip the sections in liquid photographic emulsion in a dark room under a safe-light.

X-rays were examined to judge, which slides were worth exploring further. It is important to dip sections as soon as possible after autoradiography.

The following was brought together in the darkroom:

- A small water bath with water, preset to run at 43^{o}C.
- Ilford K5, which had been stored at 4^{o}C for no less than 3 months.

- A 50 ml falcon tube containing 40-50 ml of freshly made 600 mM ammonium acetate with glycerol at 0.5% prewarmed to 43^{o}C.
- A glass rod, a dipping chamber, prewarmed in the water bath.
- A flat metal plate placed on an ice tray.
- Staining racks.
- A light-proof container, tape and scissors.

Procedure:

- Approximately 20 mls worth of the liquid photographic emulsion shreds (K5) by volume were placed into a clean 50 ml falcon tube and 'melted' in the 43^{o}C water bath, using the glass rod to ensure no bubbles/lumps occur whilst adding an equal volume of ammonium acetate solution.
- The K5/Ammonium Acetate solution was placed in the dipping chamber held in the water bath, at 43^{o}C.
- A couple of the spare slides were dipped to check the quality of the emulsion and to remove air bubbles.
- Taking the slide, with the section upper most, each slide was slowly immerse in the dipping solution twice with a smooth action ensuring an even coating.
- The underside of the slide was scrapped on the chamber and the back of the slide mopped with a paper towel to remove excess emulsion sticking to the cooled metal plate. The slide was then placed on the cooled metal plate, for approx 10 mins, with emulsion uppermost, to set.
- The dipping chamber reservoir was topped up as necessary to ensure sufficient emulsion was available for the slides to be dipped.

- Once the emulsion was set, the slides were placed in a staining rack and positioned upright in a light proof box containing silica gel for 1-2 hrs at room temperature, to dry.
- Finally, the slides were placed in a slide box with silica gel, sealing the edges, and wrapping in foil. Left at room temperature for a further 2-3 hrs prior to placing at 4^{o}C marked with the expected development date (it pays to have a few at an advanced date, in order to develop early). Development could be anything from 1-12 weeks depending on the strength of hybridization signal.

3.3.7.1. Development of emulsion-coated sections

Emulsion coated sections were removed from 4^{o}C 30 - 60 mins prior to processing, to allow slides to reach room temperature.

Meanwhile, the following was prepared in staining troughs:

- Developer. The temperature of the developer solution should be 20^{o}C $\pm$ 1^{o}C as temperature affects the speed of development (an increase of 1^{o}C adds approx. 15 secs to the developing time). This solution has a short life, and after two or three racks of slides, should be remade.
- Freshly made stop solution (see 3.2.23).
- Fix solution (see 3,2,24). (An alternatively solution is Rapid fixer HYPAM)
- Wash: single distilled/deionised water.

In the darkroom - test slides were transferred to a rack and processed as follows:

Developer:	6 mins
(longer time will give bigger grains, but also more background)	
Stop:	2 mins
Fix:	6 mins
Water:	5 mins
(keep slides moist until required)	

Only the development time is critical, the others are recommended minimum times. At this point it is prudent to check the test slides under the microscope before proceeding with the remainder.

Slides were washed in a sink with slowly running water for 30 mins - 2 hrs.

3.3.8. Counterstaining

A number of different stains may be used to counterstain sections following in-situ hybridisation. Methylene blue is quick and easy, but Hematoxylin and Eosin (H&E) is used here, resulting in the nucleus blue/purple (*hematoxylin*) and cytoplasm pink (*eosin*).

- Slide racks were removed from sink one at a time, remembering to keep slides moist at all times, rinse in double distilled water.
- Immersed in hematoxylin for 10 mins.
- Rinsed in tap water until blue (approx. 5-10 mins).
- Excess hematoxylin was removed with 1% acid ethanol until background clears and nuclei remain blue, should take only a few seconds. If nuclei remain too dark you may subsequently have problems seeing the grains against a dark

background.

- Returned to tap water until blue (approx. 5-10 mins).
- Immersed in 1% eosin for 2 mins.
- Returned to tap water until cytoplasm is pale pink and emulsion relatively clear (approx. 5-10 mins).
- Dehydrated in the following:
- 70% Ethanol for 5 mins (may be left longer if eosin stain is too strong).
- 95% Ethanol for 5 mins
- 100% Ethanol for 5 mins x 3
- Xylene for 5 mins (slides can be left here at this stage).
- Slides removed one at a time from the final xylene.
- A couple of drops of DPX mountant was applied to each section, covered with a tissue polished coverslip and any air bubbles pushed out with tweezers (if dehydration and clearing has not been completely thorough, droplets will appear under the cover slip obscuring the view of the section).

Slides were then ready for observation under the microscope.

3.3.9. Protocol for capturing images

The analysis of the in-situ hybridisation autoradiographs and dipped slides was performed at Pfizer Cambridge with the assistance of Alistair Dixon, Peter Wooding, and Janet Small, and later at the Anatomy department with the assistance of Marie Watkins and Professor Martin Johnson.

• CCD relay camera was connected to the computer with a lens. Either the MTI CCD 72 connected to a Nikon PN-11 52.5 plus / NIKKOR 105mm 1:2,8 # 244228 when

viewing from the illuminator light box, or the Leica DC 200 connected to a Leica DN RB microscope when viewing microscopically.

- M4 icon starts the image software.
- Initially the **graphic window** opens
- Using the scan area toggle ensure the following sensitivity fields are set:

Target defined in: current channel.		Target size: off
Target scan mode: normal.		Filled: off
Measures: group.		Single pixel: count
Keep only last scan shapes		
Segmentation range		highlighting
Width: levels 91	density ROD	red
Centre: 0.755	*high 0.4540*	none
	Low 2.4082	

- Using the options toggle check the following fields:

Targets defined in: channel 1(ch1) or channel 2 (ch2)

Target scan mode:

Normal

Measures:

Group

Targets

Mean size 1

Keep only last scan shapes

- In the **transform window** the following fields were used:

	Target accent	
Channel: source: 1		2: destination

Toggle the following icons

Modify:

kernal size

7 X 7

3 X 3 63 X 63

N channels:

Channel selection: all channels

Filters: target accent

Point operators: none

• In the **calibration window** and viewing a slide, the following settings applied:

Toggle the establish icon to set the distance as a X 40 lens

Calibrated

Distance in μm		cal. directions
Horizontal:	84 pixels 40μm	horizontal and vertical
Vertical:	84 pixels 40μm	

• Ensuring the microscope is visible to the camera at the slide at the top left hand side of the microscope, near the camera was positioned to 50:50 and the computer image adjusted with the fine focus of the microscope.

• The middle mouse button moves the curser from the computer screen to the image screen. From the toolbox icon a shape was selected for the area of interest. With the Ctrl key held down the mouse will increase or decrease the size of the area for scanning. With the Alt key depressed the area for scanning can be rotated.

• Ctrl plus Page Up or Page Down keys switched the image between channels.

• With the above settings, the silver grain representation was in channel 2 and the full image was in channel 1. Through selecting a defined area and confirming with the

right hand mouse button, counts were established as relative optical density (ROD). Fields ranging from ROD, scan area, total target area to count density are chosen from the options toggle in the graphic window and data stemming from this exercise was stored in the graph window for copying and pasting into excel for subsequent analysis.

- Images were copied and saved through either copying the whole image or pasting a selected area of interest onto a black background in another channel and saved.
- Subsequent manipulation of the images was carried out in Adobe photoshop and transferred to power point for printing.

3.3.10. Brief summary of *In-situ* Protocol

Tissues from embryo day stages 8.5d, 9.5d, 10.5d, 13.5d, 15.5d and 17.5dpc were prepared as described, and sections were cut, fixed and stored under ethanol. Probes were prepared as described above. *Sox4* was the only probe that failed to provide sufficient counts to proceed with hybridization, and was not pursued further. Each of these labelled probes was then incubated with sections. Slides were processed as described, put down to film for 6 days, re-exposed for a further 21 days, then dipped and left for a further 21 days, developed and counterstained.

Legends for Figures 64 and 65:

Figure 64: Sections of the autoradiographs captured with a digital camera.

Figure 65: A selection of slides, at x 6 magnification, captured on film.

Figure 64: Autoradiograph composition of slide images

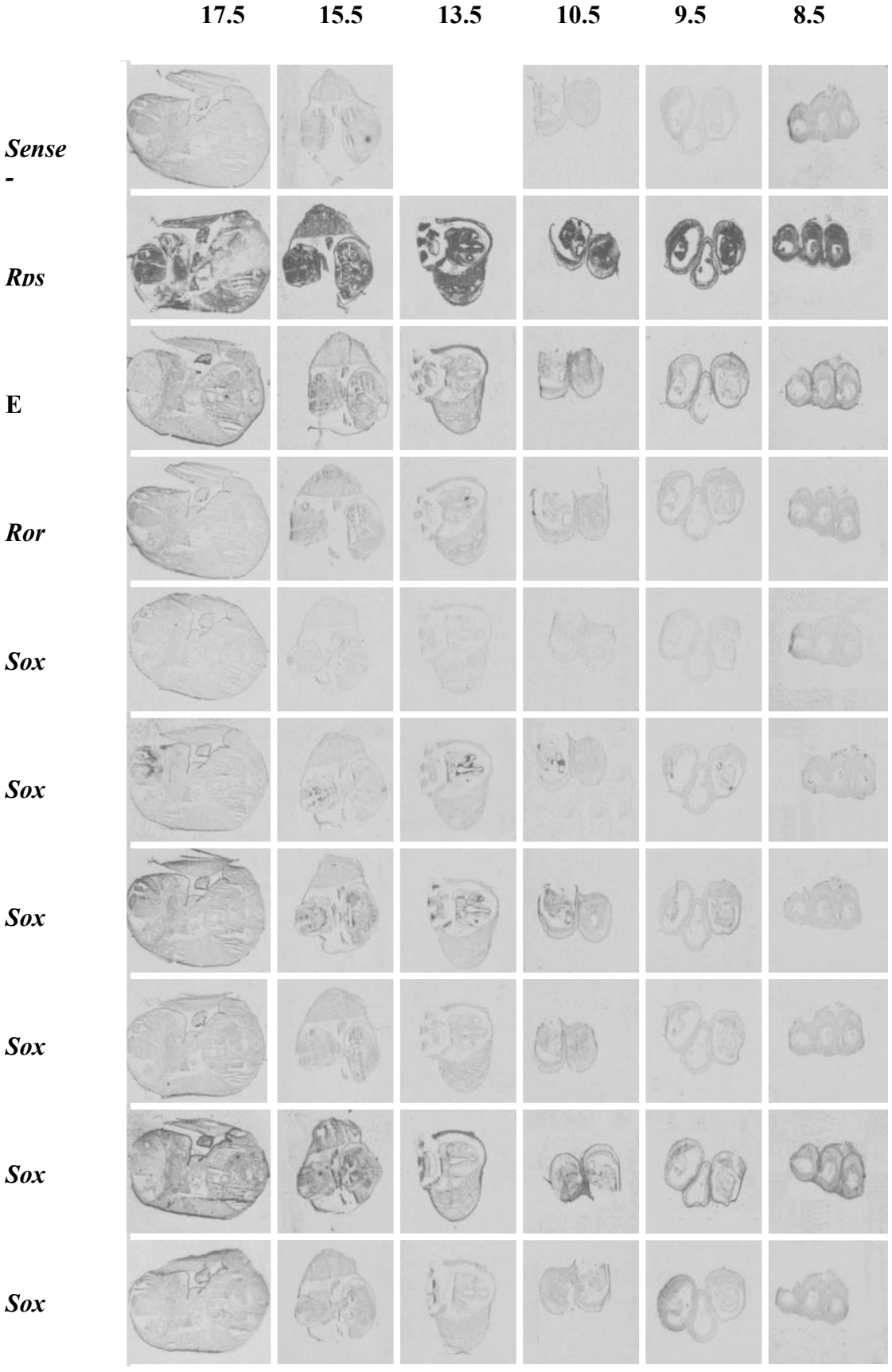
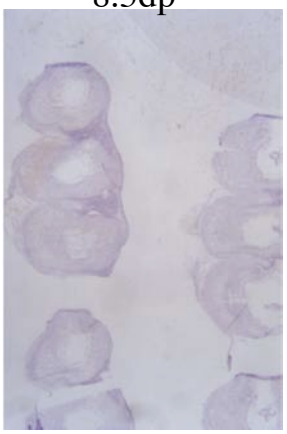


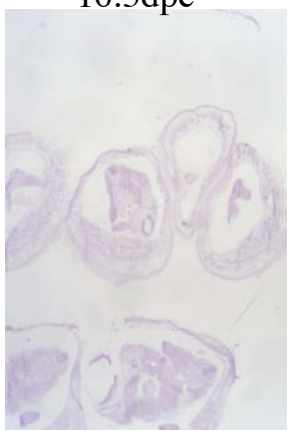
Figure 65: Images of H&E stained tissue sections,

8.5dp

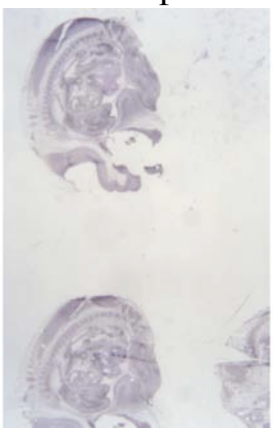


8.5dp

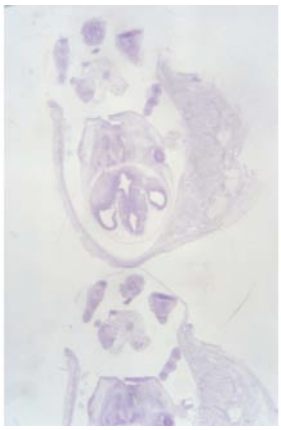
10.5dpc



12.5dp



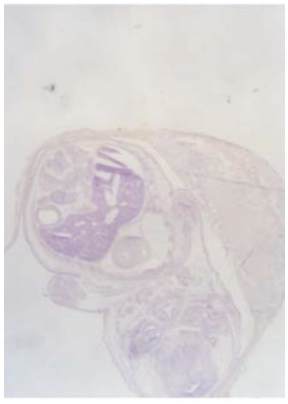
13.5dp



13.5dp

15.5dp

15.5dp



17.5dpc

3.4. Results

3.4.1. Visual Analysis of Slides

A composite of the autoradiographs of slide images is illustrated in Figure 64, showing the 10 probes used in this study on each of 6 staged foetuses. This figure gives an over view of the various probes and their broad localization. It also illustrates the importance of tissue orientation at the time of cryostat sectioning and the low resolution obtained through autoradiography. Figure 65 illustrates the different orientations of the staged embryo sections that were taken. This range of orientations was an unfortunate result of cryostat sectioning the fetuses in their embryonic sacs. Thus, some sections were sagittal (17.5d), some transverse (12.5d, 15.5d), and others of indeterminate orientation (15.5d, 13.5d, 12.5d, 10.5d, and 8.5d), the precise orientation being revealed only at the point of microsopic examination. Appendix 2 highlights the main characteristics used to determine the foetal staging.

The top row of Figure 64 shows hybridisation of the negative control sense probe (*Psmc1)*, which shows no signs of specific hybridisation, other than at the edge of the sections. The positive control probe for the housekeeping gene *Rps 29* is shown in the second row in Figure 64 and hybridises throughout all tissues of all sections, as appropriate for a housekeeping gene. Of the remaining probes in this figure *Egf*, and *Sox16* show hybridisation at the edge of the tissue sections, that appear distinct from non-specific edge effects found with the negative control. *Rora, SoxLZ, Sox15* and *Sox17* showed very little evidence of specific hybridisation, *Sox2* and *Sox6* show the most interesting hybridisation signals to internal structures.

Through dipping the slides, a closer examination of each section was made possible. The process of dipping slides in a photographic emulsion enables a more precise localisation of the radioactive hybridisation position. Microscopic

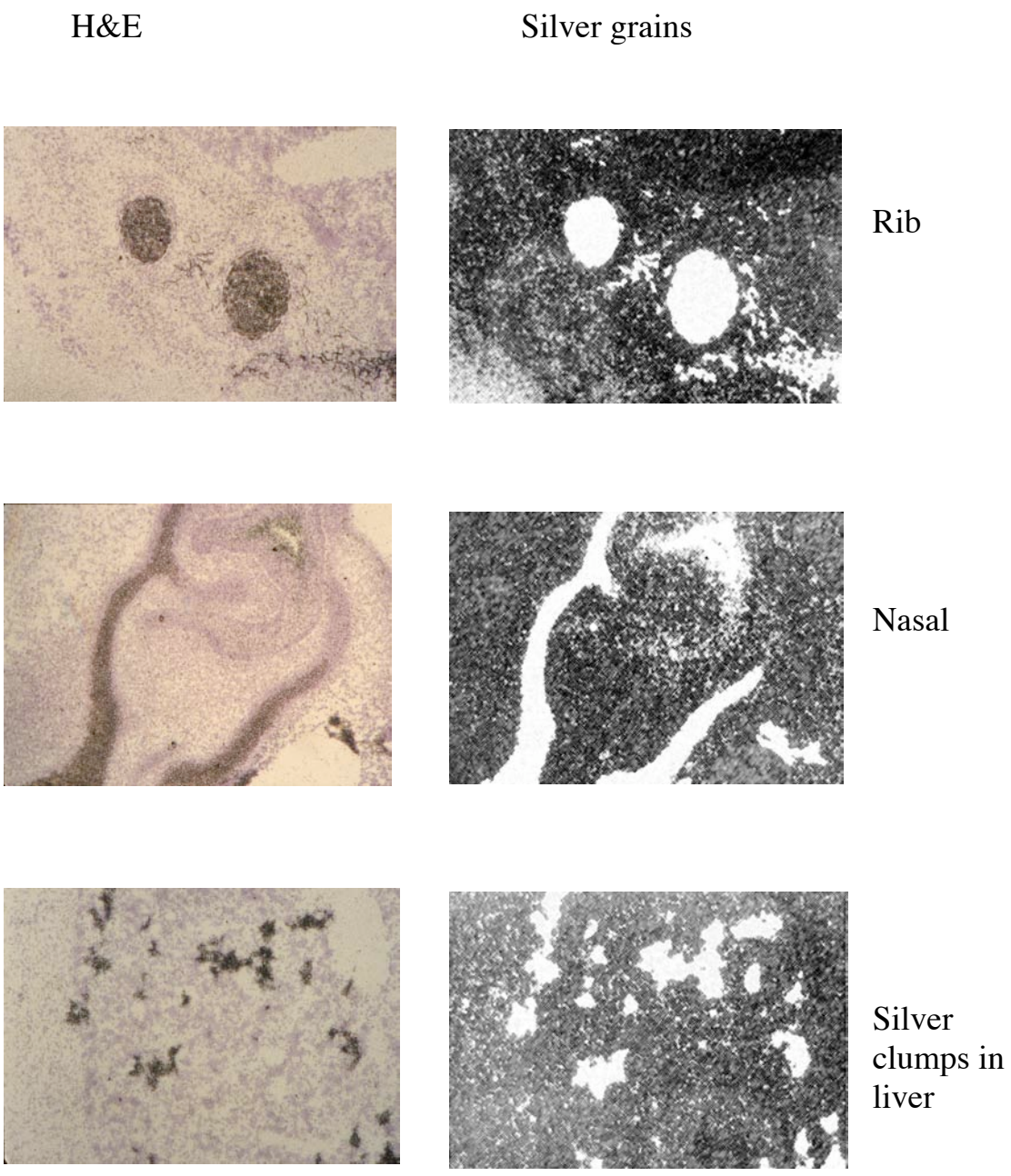
identification of silver grains under dark-field, when compared to the H&E staining assists the process of locating regions of hybridisation. The slides are coated in a gelatinous solution containing silver salts; beta particles from the S^{35}, which radiate out from the section, reduce silver in the emulsion to metallic silver (halide). On developing the sections the sensitized silver ions show as black specks under a light microscope and reflect light in the dark-field. It is this dark-field image that is most informative for data analysis. All resulting figures of *in-situ* counterstaining and silver grains are at x10 magnification with the exception of Figure 75, which is at x40 magnification.

Results from dipped sections for *Egf, Rora, SoxLZ, Sox15, Sox16* and *Sox17* showed few regions suggesting specific *in-situ* signals and those regions were faint in comparison to *Rps29, Sox2* and *Sox6.* Because the object of this exercise was to compare *in-situ* results with those from RTPCR, it was decided to focus on the *Sox2* and *Sox6* data from this series of sections. The results were also compared with those from published papers on gene expression, which were more extensive for these two genes. There was very little literature for the *Sox* genes LZ, 15, 16 & 17, all of which gave weak *in-situ* signals .

The *Egf* probe hybridised to the outer embryonic sac of the early stages (8.5-12.5dpc), placental tissues of 15.5dpc, yolk sac of 17.5dpc, with binding to few other identifiable tissues (see eye below), and gave a clumping effect in the liver (data not shown). This latter effect was regarded as an artefact, since the sense probe, *Psmc1,* also gave this sort of signal. This negative control probe highlighted a number of other potential artefacts, illustrated in Figures 66 and 67, notably an edge effect to some sections, and autofluorescence from cartilage and retina pigment layer. These

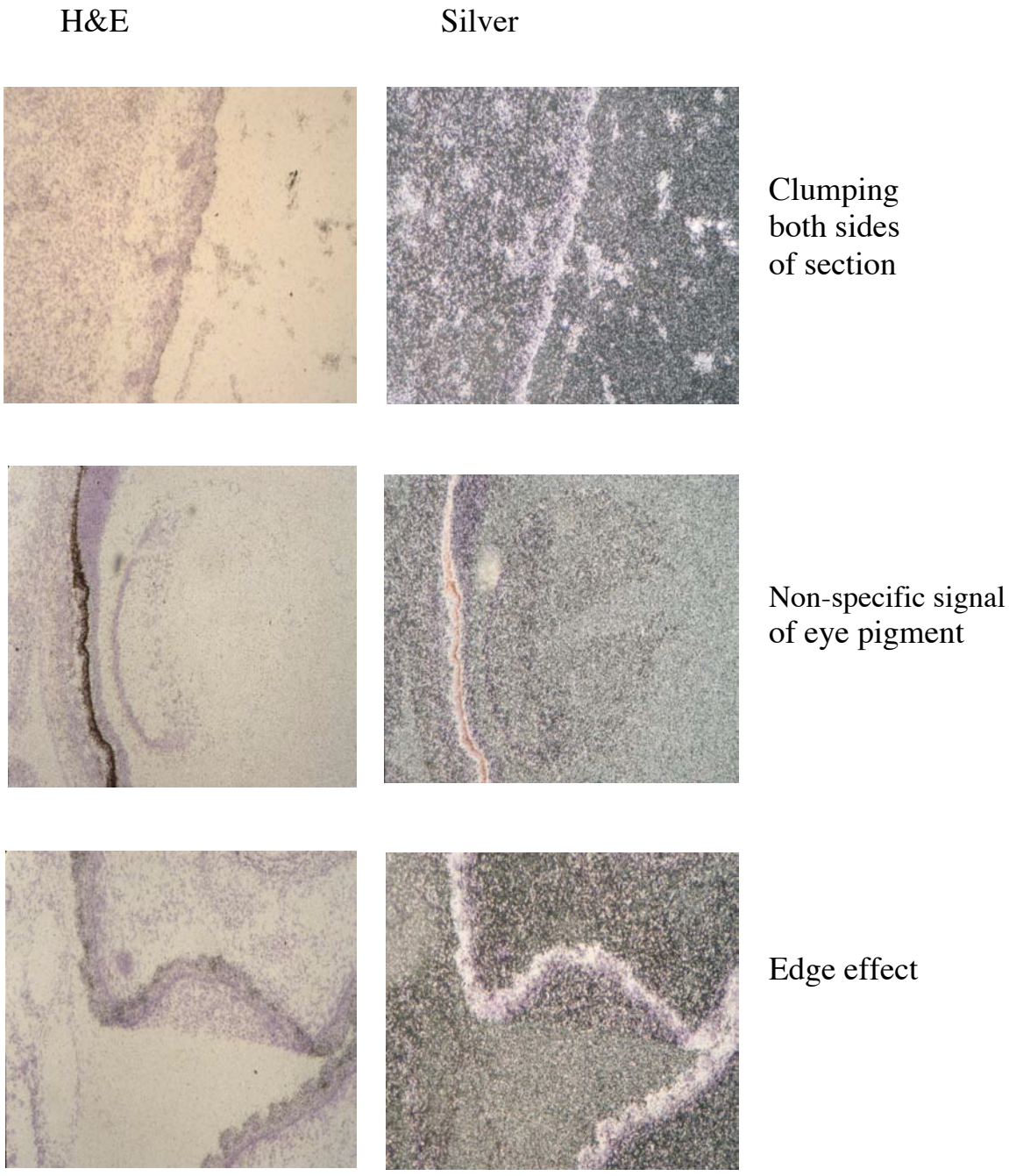
artefactual effects were ignored as non-specific signals when found during the charting of expression profiles using the two more successful probes *Sox2* and *Sox6*. Specific hybridisation for these two probes are set out as tables in Figures 68 and 69.

Figure 66: Artifacts of Silver Staining.



Microscope images at x 10 magnification, from H&E staining and silver dipping, captured on film, illustrating non-specific signal.

Figure 67: Silver Staining Artifacts



Microscope images at x 10 magnification, from H&E staining and silver dipping, captured on film, illustrating non-specific signal.

Figure 68: Chart of *Sox2 In-situ* hybridisation

Tissue/day	Forebrain/Head region	Midbrain	Hindbrain	Spinal cord	Heart	Lung	Liver	Intestine
17.5	Telencephalon Sensory layer of retina	Dorsal thalamus	Mescencephalon	Ependymal layer of spinal canal	ND	Lining cells of oropharynx	faint	ND
15.5	Telencephalon Cells lining the nasal passage Sensory layer of retina Optic chiasma	Dorsal thalamus Olfactory epithelium Recess of 3rd ventricle	Mescencephalon	Ependymal layer of spinal canal	NA	Lining of mouth, Olfactory epithelium. (Rathke's pocket) Bronchi	ND	ND
13.5	Semilunar ganglion Cochlea	Dorsal thalamus	Mescencephalon	Ependymal layer of spinal canal	NA	NA	ND	NA
12.5	Telencephalon	Present	Myelencephalon	Spinal canal	NA	NA	ND	ND
10.5	Ventricle edge of Telencephalon	ND	Myelencephalon	NA	NA	NA	NA	NA
8.5	Forebrain Stomodaeum region	NA	NA	Neural fold	NA	NA	NA	NA
Tissue/day	**Oesophagus/ stomach**	**Urogenital**	**Kidney**	**Gonads**	**Bladder**	**Forelimb/Rest of body**	**Placenta**	**Yolk sac**
17.5	Duodenum	ND	Kidney	NA	ND	Epidermis Cartilage	ND	NA
15.5	Mucosal lining of stomach	Lining of urogenital sinus	ND	NA	NA	Epidermis	NA	NA
13.5	Oesophagus Mucosal lining of stomach	ND	ND	NA	NA	Epidermis	ND	ND
12.5	Stomach	ND	ND	NA	NA	Epidermis	ND	ND
10.5	NA	NA	NA	NA	NA	NA	ND	ND
8.5	NA	NA	NA	NA	NA	NA	ND	ND

ND – signal Not Detected
NA - tissue Not Available for informative diagnosis

Figure 69: Chart of *Sox6 In-situ* hybridisation

Tissue/day	Forebrain/Head region	Midbrain	Hindbrain	Spinal cord	Heart	Lung	Liver	Intestine
17.5	Sensory layer of retina Ganglion	ND	ND	ND	ND	ND	Present	ND
15.5	Nasal capsule. Cortex Meckel's cartilage (jaw)	Present	ND	Ependymal layer of spinal canal Spinal ganglion.	NA	Trachea	Present	ND
13.5	Telencephalon Nasal Cartilage Meckel's cartilage	ND	ND	Ependymal layer of spinal canal	NA	Bronchi	Present	NA
12.5	Neopallial cortex Nasal Cartilage Meckel's cartilage	Midbrain Striatum	Medulla	ND	NA	NA	ND	ND
10.5	Telencephalon Neural fold	Myelencephalon	ND	NA	NA	NA	NA	NA
8.5	Squamous endoderm	NA	NA	Neural fold	NA	NA	NA	NA
Tissue/day	**Oesophagus/ stomach**	**Urogenital**	**Kidney**	**Gonads**	**Bladder**	**Forelimb/Rest of body**	**Placenta**	**Yolk sac**
17.5	ND	ND	ND	NA	ND	Cartilage	ND	NA
15.5	Duodenum	ND	ND	NA	NA	Cartilage	NA	NA
13.5	ND	ND	ND	NA	NA	ND	ND	ND
12.5	ND	ND	ND	NA	NA	NA	ND	ND
10.5	NA	NA	NA	NA	NA	NA	Present	ND
8.5	NA	NA	NA	NA	NA	NA	ND	ND

ND – signal Not Detected
NA - tissue Not Available for informative diagnosis

Of the 90 cDNAs represented in the mouse foetal panel, those not represented in the *in-situ* analysis are the 5 controls – two glycogen blanks, genomic mouse, rat and human DNA, 11 adult tissues, a single 9.5dpc and two 11.5dpc samples. Of the remaining 71 tissue samples, only 50 tissues were presented for analysis, those missing were not identifiable in sections available (as shown by "NA" in the charts). For all those tissues that were identified in sections and so available for analysis, the concordance between RTPCR and *in-situ* analysis was expressed in percentage terms for each of the two *Sox* oligo probes analysed here.

For *Sox2*, there were 31 cases in which a signal was detected in both the *in-situ* and the RTPCR tissue samples, and 4 cases in which no signal (ND) coincided with a negative value for RTPCR. Collectively, this gave a 70% score where the RTPCR signal was confirmed by the *in-situ* signal.

The remaining 30% (15) of cases were all where there was no signal (ND) detected *in-situ,* but where a positive value had been found by RTPCR. There were no cases of tissues showing a signal from *in-situ* analysis, with a negative signal in the RTPCR. Reported expression of *Sox2* primarily concerns the preimplantation stages [1], developing CNS [2], and the lens [3] [4]. Precise sub-localisation of *Sox2,* expression was found here for the following areas: the ganglion, amacrine, synaptic and plex layers of the eye (17.5dpc), cells lining the oropharynx (17.5dpc), Rathke's pocket (15.5dpc), lining of the urogenital sinus (15.5dpc), the vibrissae (15.5dpc), the bronchi of the lung (15.5dpc), the urethra (13.5dpc), and the ependymal layer of the spinal cord (13.5 and 15.5dpc), and much of the developing brain and spinal cord from 8.5dpc through to 17.5dpc, confirming it's importance as a lens and CNS developmentally important gene.

For *Sox6*, there were 20 cases where there was signal for *Sox6* in both the *in-situ* and the RTPCR tissue samples, and 2 incidences where no signal (ND) was coincident with a negative value for RTPCR. Collectively, this gave a 56% score where the RTPCR signal was confirmed by the *in-situ* signal. The remaining 44% (28) of cases were again where no signal (ND) was found in the *in-situ* analysis, but where a positive signal had been found by RTPCR. Again, there were no cases of tissues showing signal from *in-situ* with a negative signal in the RTPCR. The majority of the known literature relating to *Sox6* expression reinforces the view of this gene as an essential component of the developing CNS and cartilage formation [5-9]. Confirmation of this gene's involvement in cartilage formation is shown here through hybridisation to cartilage structures in the forelimb, ribs, Meckel's and nasal cartilage. It was also found in the 10.5 dpc placenta, 15.5 dpc duodenum, the liver from 13.5 – 17.5 dpc, trachea of 15.5 dpc and bronchi of lung at 13.5 dpc indicating that it may have a wider role in development.

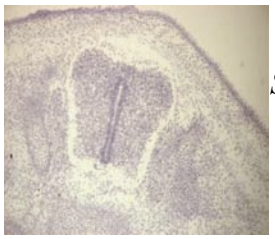
One of the most striking regions of expression was the spinal cord region in 13.5dpc sections (Figure 70 for 10 of the probes demonstrates the benefit of silver dipping). These regions of silver grain staining have been quantified using the method described in 3.3.9 and shown below graphically as Figure 80. A closer image of the *Sox2* probe hybridisation to the spinal cord is shown as Figure 71, which identifies the target area more precisely as the ependymal layer of the spinal cord as shown by the diagram in Figure 72 (p.269, The Mouse It's Reproduction and Development by Roberts Rugh).

The top four panels of Figure 73 illustrate a region of the lung in 15.5dpc mouse. *Sox2* is compared to the control sense probe (*Psmc1*) showing specific labelling of the bronchi by the probe for *Sox2*, illustrating that the *in-situ* technique

provides useful localisation information undetected by RTPCR alone. The lower four panels of Figure 73 shows the snout region of the 15.5dpc mouse, illustrating specific hybridisation of *Sox2* to the primordial vibrissae follicles (whiskers growth site), again compared to the sense probe.

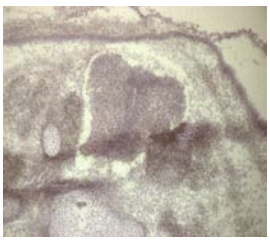
Figure 74 is of the olfactory region of the 15.5dpc mouse, showing *Sox2, SoxLZ, Sox6 and Sox16* probe hybridisations, together with the H&E counterstaining to identify the region of interest. This figure shows a strong signal for the *Sox2* probe, *SoxLZ* shows no specific signal, *Sox6* and *Sox16* are showing signal at the outer perimeter of the olfactory region, most likely cartilage. In comparison, the cartilage artefactual signal illustrated in Figure 66 when viewed under the microscope is a far brighter and intense signal.

Figure 70: 13.5d region of spinal cord hybridisation results

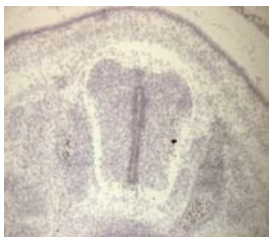


Sense - Psmc1

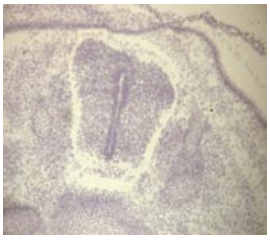
Sox 2



Rps2



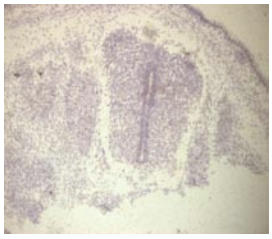
Sox 6



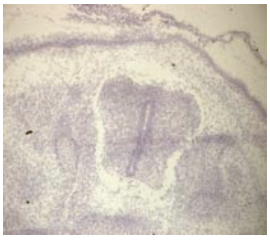
Egf

Sox 15

Rora



Sox 16

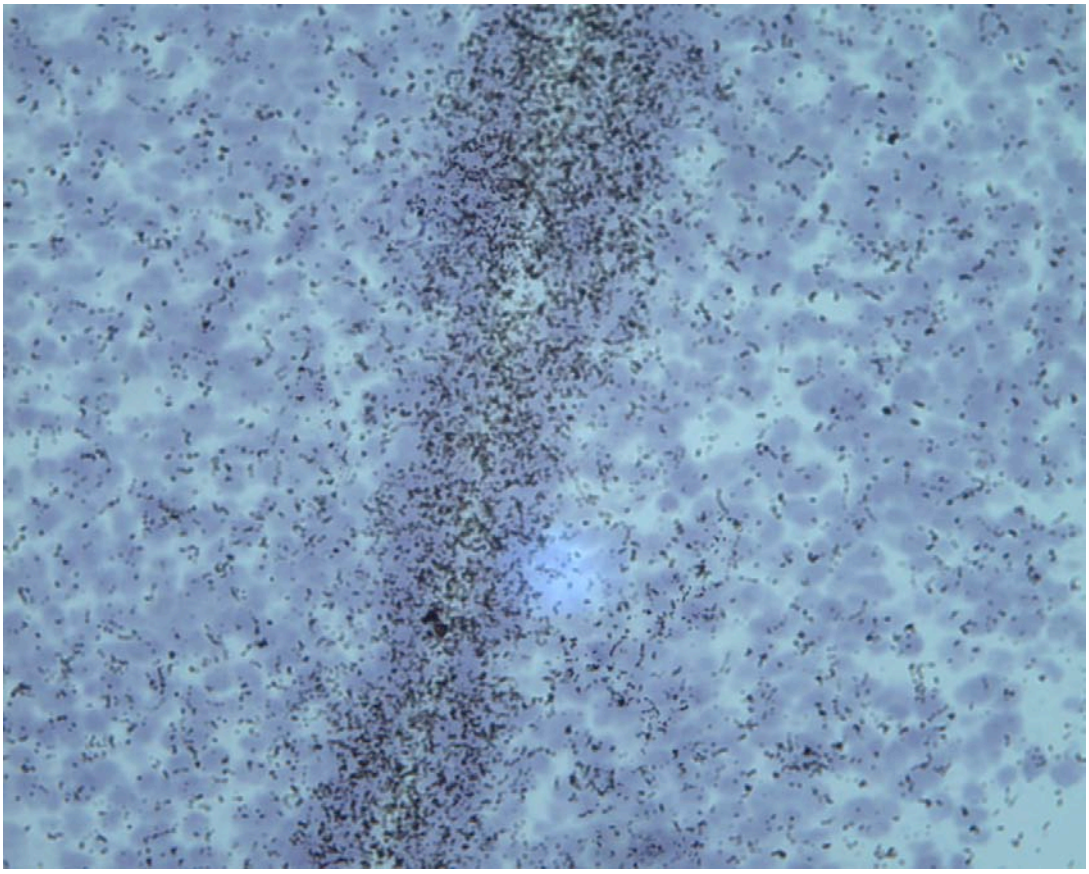


SoxLZ

Sox 17

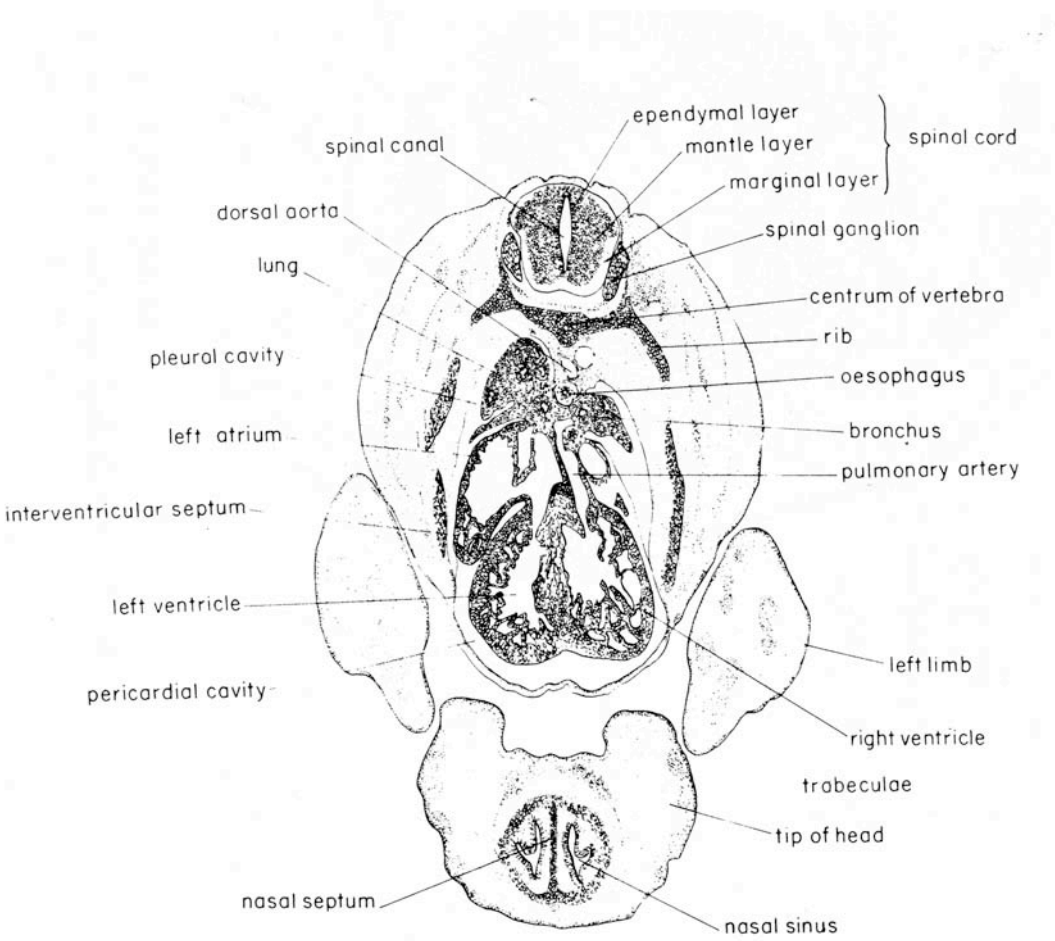
Spinal cord view of 13.5d sections, silver dipped, x 10 magnification. Illustrating the ependymal layer with all 10 probes.

Figure 71: Ependymal Layer of 13.5dpc spinal cord, hybridised to *Sox2* (x40)



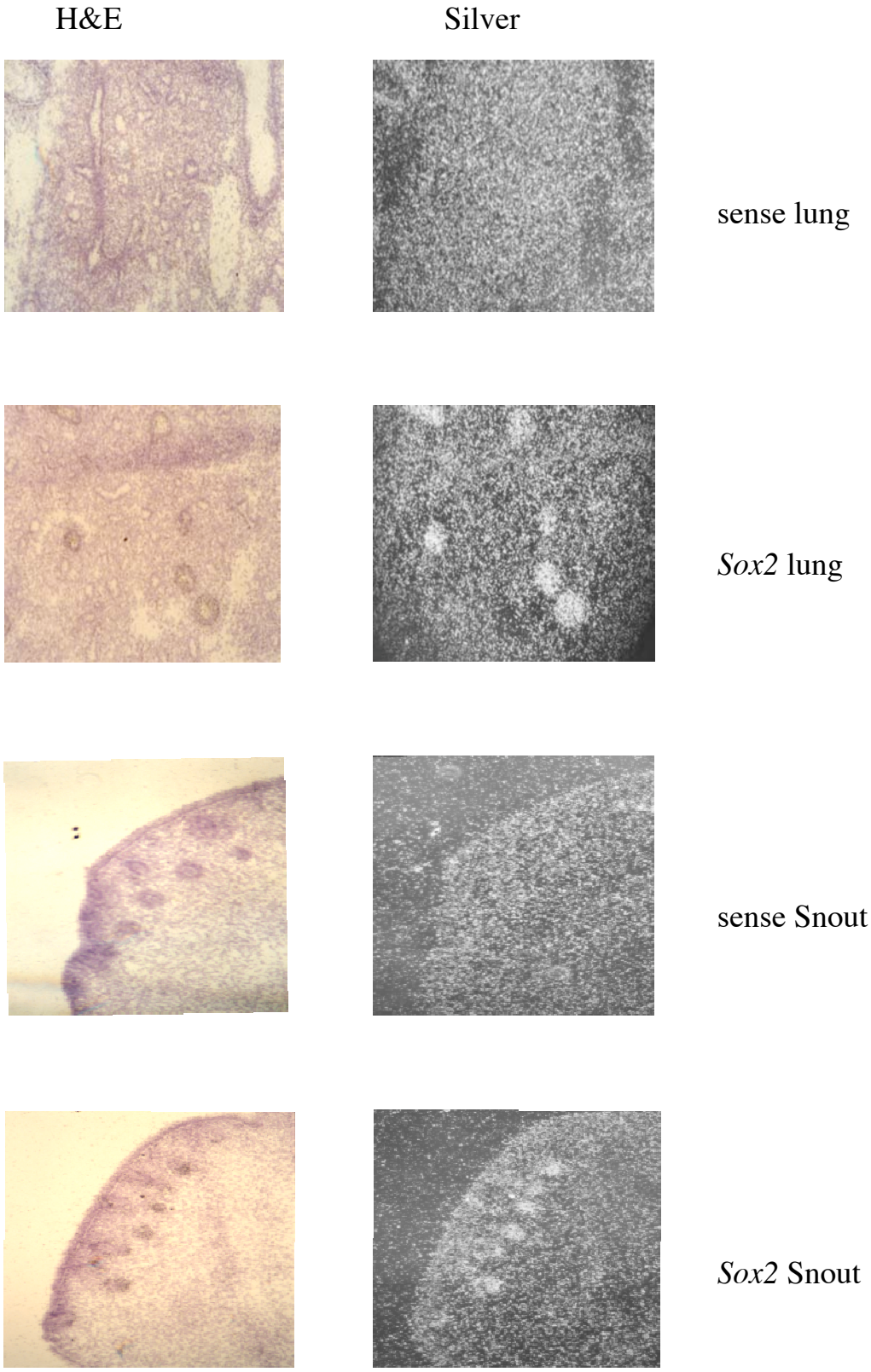
Spinal cord view of 13.5d sections, silver dipped, x 40 magnification. Illustrating the ependymal layer for the *Sox2* probe.

Figure 72: Illustration of a transverse section of a mouse embryo at 14 days gestation



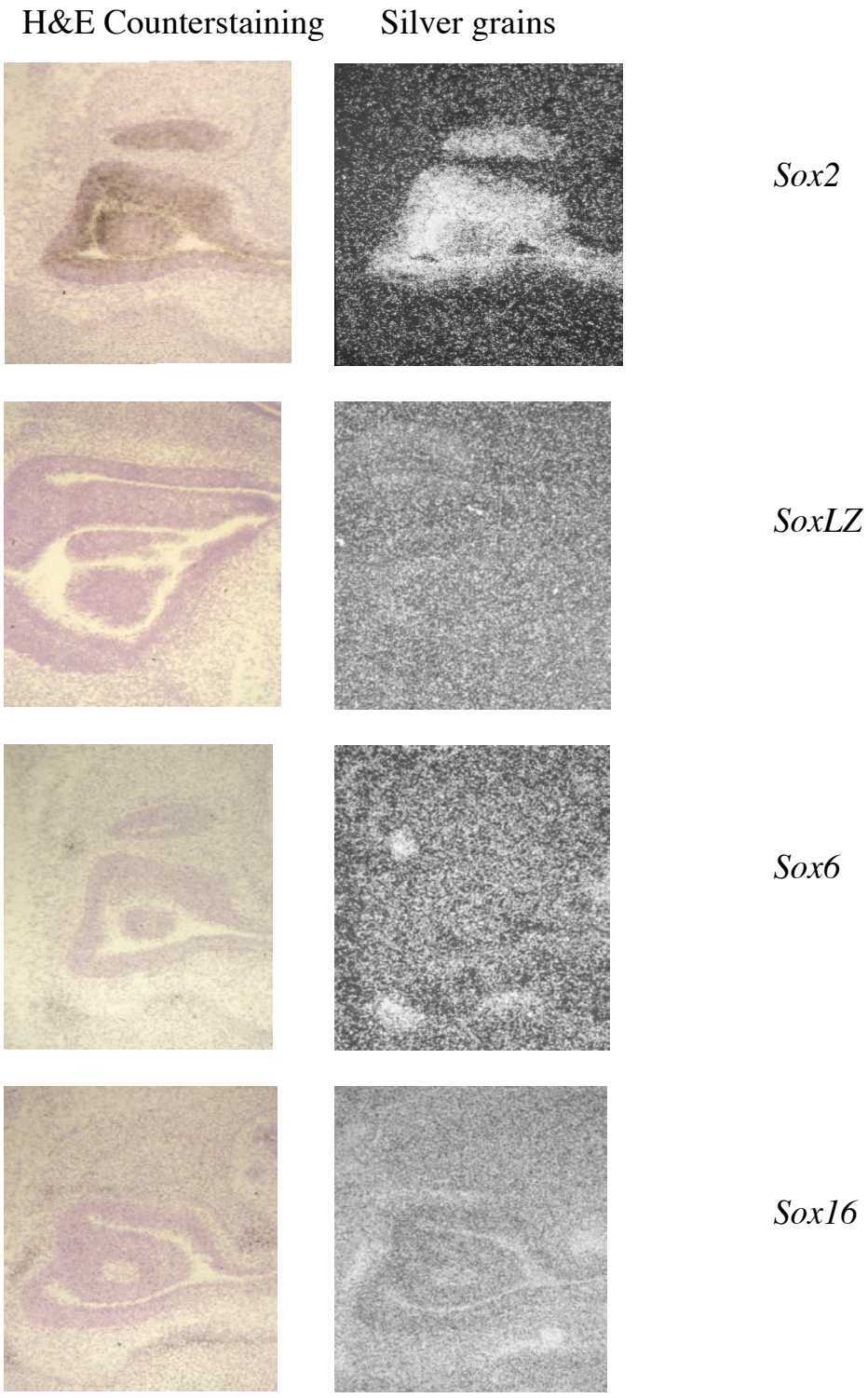
Reference: The Mouse It's Reproduction and Development. Roberts Rugh

Figure 73: Lung and Snout regions



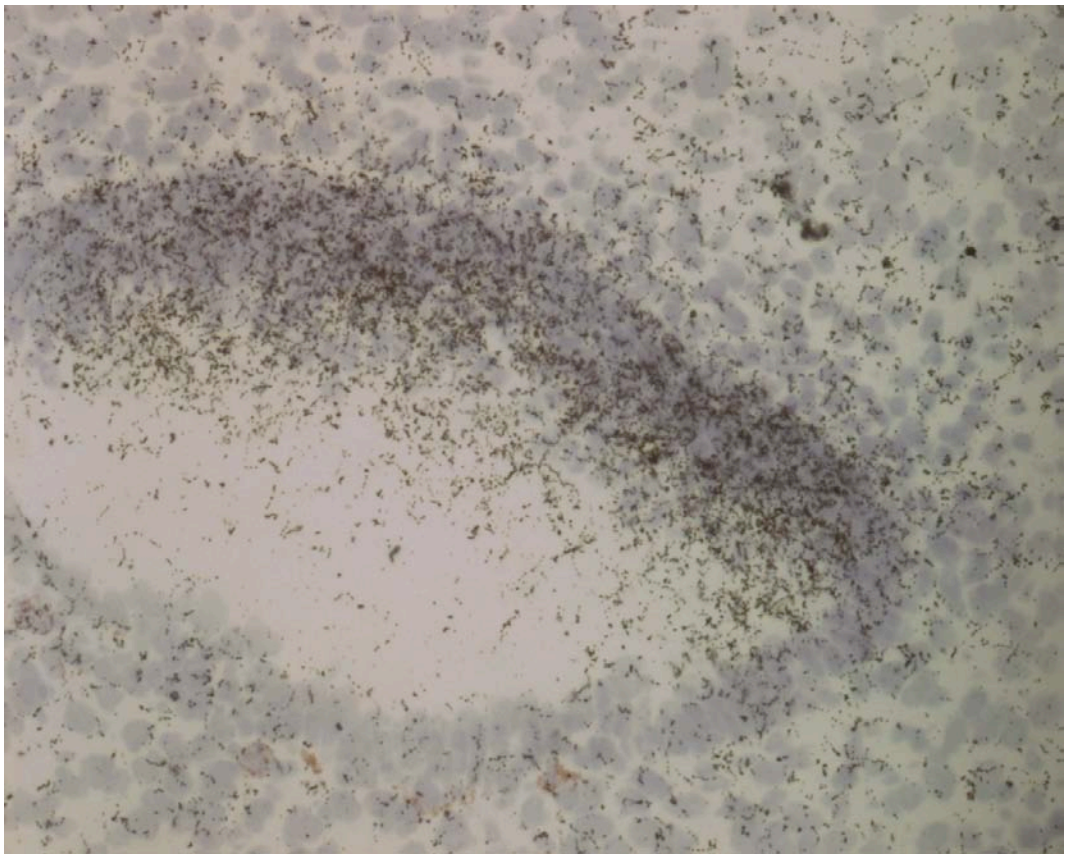
Microscope images at x 10 magnification, from H&E staining and silver dipping, captured on film, illustrating signal due to specific hybridisation.

Figure 74: Olfactory region of stage 15.5d



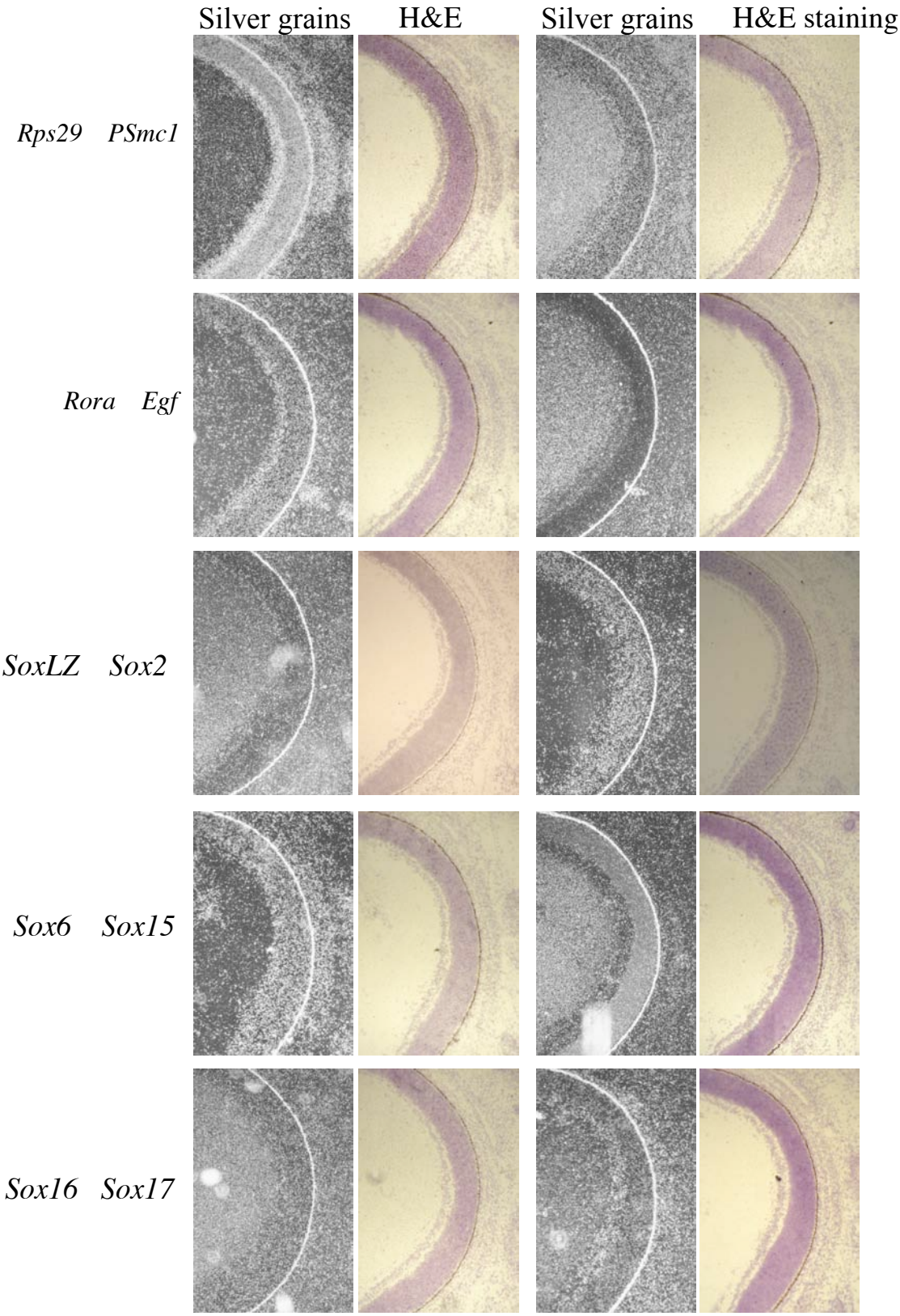
Microscope images at x 10 magnification, from H&E staining and silver dipping, captured on film, illustrating signal due to specific hybridisation in the olfactory region.

Figure 75: 15.5d stage cochlea, hybridised to *Sox2*



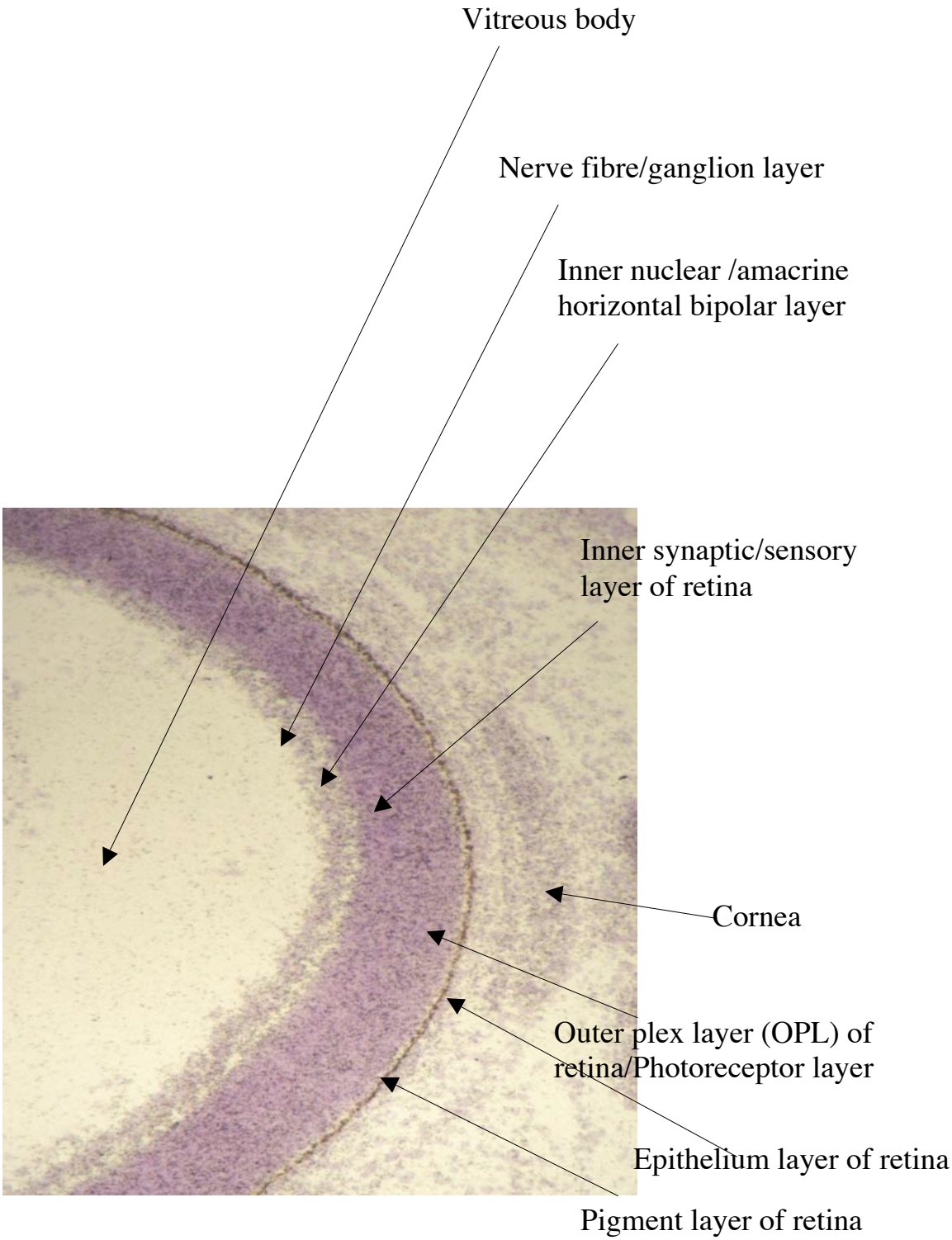
Microscope images at x 10 magnification, from H&E staining and silver dipping, captured on film, illustrating signal due to specific hybridisation in the cochlea.

Figure 76: 17.5d Eye hybridization results



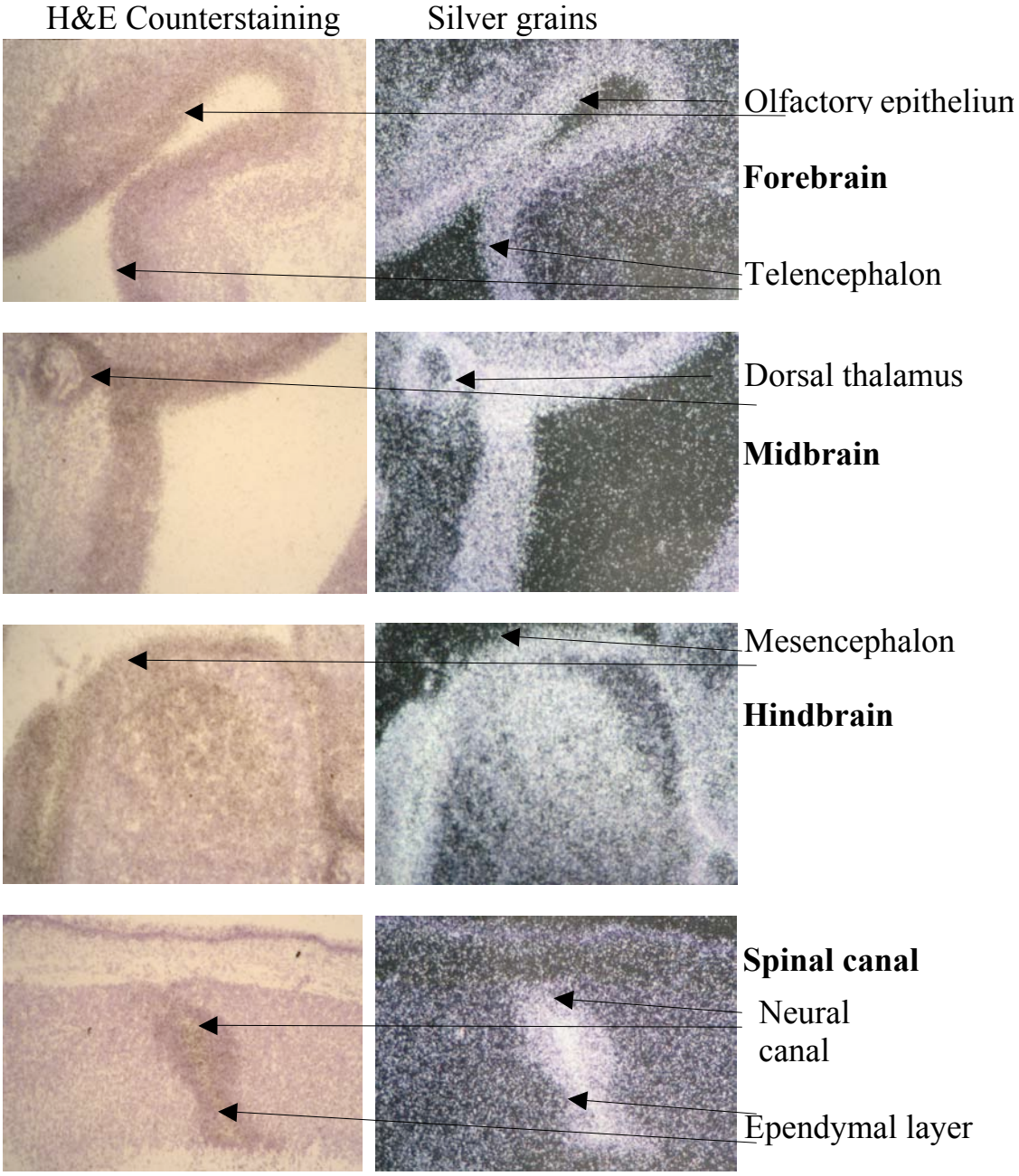
Microscope images x 10 magnification, from H&E staining and silver dipping, captured on film, illustrating signal due to specific hybridisation in the 17.5d eye. The signal in pigmented epithelium is a non-specific artefact.

Figure 77: 17.5d Eye - identification of layers



Microscope image x 10 magnification, from H&E staining, captured on film, illustrating layers of the 17.5d eye.

Figure 79: Neural tissues of 15.5d hybridised to *Sox2*



Microscope images x 10 magnification, from H&E staining and silver dipping, captured on film, illustrating signal due to specific *Sox2* hybridisation in the 15.5d brain.

Figure 75 shows the *Sox2* probe hybridising to the 15.5dpc cochlea.

Figure 76, is a composite of H&E and DF (dark-field) images from the 17.5dpc mouse eye for all probes, and illustrates the variety of tissue layers as the embryo progresses to full term and the probe hybridisation at this stage. Figure 77 illustrates the 17.5dpc eye with the various cell types identified. These data are tabulated as Figure 80.

Figure 78: Table of specific silver staining in the eye

Probe /tissue	Vitreous body	Nervefibre /ganglion layer	Inner Amarcrine /nuclear layer	Synaptic /sensory layer	Outer plex. /nuclear layer
Rps29	-	+	+	+	+
Psmc1	+	-	+	-	-
Rora	-	+	+	-	-
Egf	+	-	+	-	-
SoxLZ	+	-	+	-	-
Sox2	-	+	+	+	+
Sox6	-	+	+	+	+
Sox15	+	-	-	+	+
Sox16	+	-	-	-	-
Sox17	-	+	+	-	-

From the tabulated results it can be seen that *Egf* and *SoxLZ* have similar expression patterns, as do *Sox2* and *Sox6*. It is known from studies in chick that *Sox2* is an important gene in eye development [10], but the co-expression of *Sox6* for this stage has not to date been recorded. The genes *Rps29, Psmc1, Rora, Egf, Sox2, Sox6, and Sox15* (but not *Sox16, Sox17 or SoxLZ)* have all been reported previously as being found in mouse adult eye [11], though the precise location had not been identified.

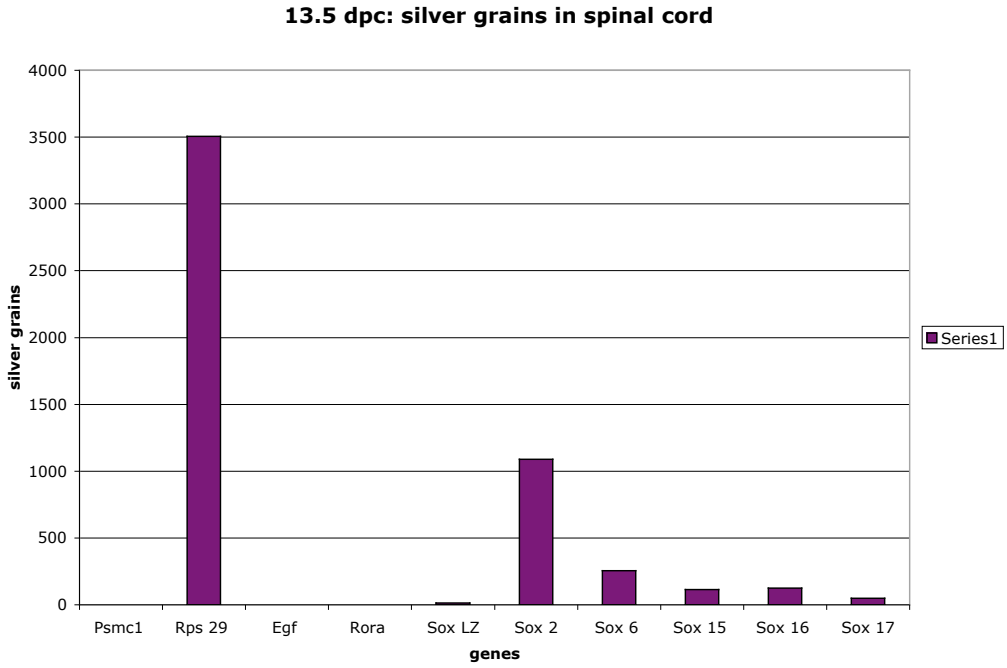
Figure 79 is a composite of various brain regions showing *Sox2* localisation at 15.5dpc. This study confirms the presence of *Sox2* expression in brain and spinal cord [12] for the stages 13.5dpc through to 17.5dpc and begins to provide some anatomical sub-localisation.

3.4.2. Digital Analysis of Slides

The ependymal layer of the spinal cord for 13.5d sections was captured digitally from the microscope and analysed using an MCID image M4 analyser software. This software measures the intensity of silver grains in a user-defined region. The output file contains the scanned area illustrating the total region analysed and total target area as a measure of silver grains present in the scanned area.

Figure 80: Data and Graph of 13.5d ependymal layer silver grain counts

Gene	Scanned area	Total target area
sense	44457	0
Rps 29	12042	3507
Egf	20148	0
Nuc. hor.rec	20148	0
Sox LZ	43765	13
Sox 2	48082	1088
Sox 6	38006	254
Sox 15	23165	115
Sox 16	31722	124
Sox 17	44259	50



The graph illustrates the silver grain concentrations as a measure of probe intensity for the ependymal layer of the spinal cord in Figure 70. The *Sox*2 probe hybridised significantly more to this region than any other probe with the exception of *Rps29,* confirming the visual analysis. The antisense control gene *Psmc1*, showed no recordable signal. *Egf*, which had shown a trace signal at 35 cycles in the adult spinal cord [13], gave no measurable signal in the ependymal layer for the 13.5 dpc spinal cord in this study. *Rora*, had shown a weak to moderate signal at RTPCR in the adult mouse panel [13], but similarly gave no measurable signal in the ependymal layer for the 13.5 dpc spinal cord in this study. As these genes were not included at the RTPCR step for fetal stages, the lack of an *in-situ* signal may be due to these genes not expressing in foetal spinal cord or merely not being present in the ependymal layer. A more thorough study would include these genes at the RTPCR stage for confirmation.

However, for the Sox gene family, RTPCR was undertaken using the fetal panel, and so the results from each approach can be compared. The strong RTPCR signals for *Rps29* and *Sox2* found in the 13.5 dpc spinal cord is localised here to the ependymal layer. In contrast, the strong RTPCR signal for *SoxLZ* is not mirrored in the *in-situ* hybridisation for the ependymal layer, nor is there a signal in tissues surrounding this layer, which may reflect the poor overall signal from this probe at the level of *in-situ* or may be because the region of the spinal cord expressing *SoxLZ* is not represented in this analysis. Both *Sox2* and *Sox6,* gave a strong RTPCR signal at all spinal cord stages; in this analysis *Sox6* is expressed at one quarter of the *Sox2* level in the region of the ependymal layer shown, illustrating a level of quantitation between genes, which can not be deduced from the RTPCR profiles alone. The probes for *Sox15, Sox16* and *Sox17* show 10.5%, 11.5% and 4.5% respectively of the signal represented by the *Sox2* probe, for the ependymal layer of the spinal cord at 13.5 dpc,. The RTPCR profiling for these genes gave trace amounts for the whole spinal cord from *Sox15* and *Sox16* and weak to moderate amounts of signal for *Sox17.* The levels of expression found by RTPCR relate to expression levels within the whole panel for a given gene but for the absolute levels to be in relationship to other genes, the method of *in-situ* hybridisation provides a clearer picture.

Measurements of this nature can be employed to analyse expression over time in a given region, as the software compensates for area differences and can be used to show differences in silver grain intensities for dissimilar sized areas. This type of analysis can be useful in quantifying expression rates in sections, once slide-dipping techniques have been mastered. Common artefacts associated with this technique include (i) high backgrounds due to poor slide handling, storage and poor darkroom facilities, (ii) edge effects as a result of rapid emulsion drying, (iii) uneven emulsion

layering over sections resulting from unmelted emulsion at time of dipping, (iv) loss of silver grains during staining, due to use of alcohol at too low a pH (pH should be 5.2) and/or immersion for too long, and finally (v) talc from gloves dropped in the emulsion can form bright centres on tissue sections, clouding genuine hybridisation data. From my experience in this study, this technique requires careful attention to detail.

3.4.3. Concluding Remarks

Data from the *in-situ* technique can be analysed in a number of different ways.

1. Visually to get a qualitative picture of expression as long as the specificity of hybridisation is secure. Hence rigorous controls and strict attention to methodology is important.

2. Quantitative information from good quality autoradiographs with ^{14}C microscales included as controls. This image can be scanned and the specific region of expression quantified computationally by measuring a defined region in relationship to the control microscales, using the NIH image software freely available from http://rsb.info.nih.gov/nih-image/, or Scion Image software. The optical density of the captured image is compared to the microscale values, resulting in a measure as nanoCuries per gram of tissue weight. This technique gives a numerical relationship to the image, illustrating the different intensities of hybridisation, between sections on the same autorad, or between different autoradographs containing the same set of microscales.

3. Sub localisation and refined anatomical information can be observed microscopically and silver grains counted (see section 3.4.2.) for a more detailed

numerical analysis. Where a refinement of the data is required, silver slide dipping permits closer examination of the defined region to identify the precise position of hybridisation. When the silver grains have been applied as fine particles and their number counted, resulting data give a clearer indication of where the individual probes have hybridised the most.

Through conducting *in-situ* hybridisation, a more detailed knowledge has been gained for some of the probes with confirmatory evidence of expression previously found at the level of RTPCR.

Discrepancies that exist between RTPCR and the *in-situ* results may reflect the lower sensitivity of the *in-situ* technique, or possibly, that a minority sub-component of the tissue was responsible for the PCR signal but proportionately too small to be detected by *in-situ* analysis. It is also possible that the correctly orientated section was not available for analysis by *in-situ*, and a wider ranged *in-situ* study could shed light on more of possible sub-localisation of the expression found at the level of RTPCR. Where an *in-situ* signal was detected, this did allow the localisation of mRNA within composite complex tissue samples used for RTPCR analysis, as illustrated in the charts for *Sox2* and *Sox6* and figures showing defined target areas. Therefore, the two approaches were complementary and additive. Overall, results shown here for the *Sox2* and *Sox6* probes provide evidence that the panel can produce reliable information.

From the diagram (Figure 3, Introduction) illustrating the structure of the *Sox* genes, the two genes *Sox2* and Sox6 are from two different groups B1 and D. They have very different structures and are thought to have different functions in the cell. The B1 group members are regarded as transcription activators [14], involved in

neurogenesis, where as the group D are involved in chondrogenesis and spermogenosis.

References

1. Avilion, A.A., et al., *Multipotent cell lineages in early mouse development depend on SOX2 function.* Genes & Development, 2003. **17**(1): p. 126-140.
2. Wood, H. and V. Episkopou, *Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages.* Mechanisms of Development, 1999. **86**: p. 197-201.
3. Kamachi, Y., et al., *Involvement of Sox1, 2 and 3 in the early and subsequent molecular events of lens induction.* Development, 1998. **125**(13): p. 2521-2532.
4. Muta, M., et al., *Distinct roles of SOX2, Pax6 and Maf transcription factors in the regulation of lens-specific delta 1-crystallin enhancer.* GENES TO CELLS Japan Osaka Univ, Grad Sch Frontier Biosci, Dev Biol Lab, Osaka, Japan, 2002. **7**(8): p. 791-805.
5. Connor, F., et al., *The Sry-related HMG box-containing gene Sox6 is expressed in the adult testis and developing nervous system of the mouse.* Nucleic Acids Res, 1995. **23**(17): p. 3365-72.
6. Lefebvre, V., P. Li, and B. deCrombrugghe, *A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene.* Embo Journal, 1998. **17**(19): p. 5718-5733.
7. Smits, P., et al., *The Transcription Factors L-Sox5 and Sox6 Are Essential for Cartilage Formation.* Dev. Cell, 2001. **1**: p. 277-290.
8. Smits, P. and V. Lefebvre, *Sox5 and Sox6 are required for notochord extracellular matrix sheath formation, notochord cell survival and development of the nucleus pulposus of intervertebral discs.* Development, 2003. **130**(6): p. 1135-48.

9. Chimal-Monroy, J., et al., *Analysis of the molecular cascade responsible for mesodermal limb chondrogenesis: sox genes and BMP signaling.* Dev. Biol, 2003. **257**(2): p. 292-301.

10. Kamachi, Y., et al., *Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development.* GENES & DEVELOPMENT Osaka Univ, Inst Mol & Cellular Biol, Osaka 5650871, Japan, 2001. **15**(10): p. 1272-1286.

11. Ringwald M, et al., *The mouse gene expression database.* Nucleic Acids Research, 2001. **29**: p. 98 - 101.

12. Zappone, M.V., et al., *Sox2 regulatory sequences direct expression of a beta-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells.* Development, 2000. **127**(11): p. 2367-2382.

13. Freeman, T.C., et al., *Expression Mapping of Mouse Genes.* MGI Direct Data Submission, 1998.

14. Bowles, J., G. Schepers, and P. Koopman, *Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators.* Developmental Biology, 2000. **227**(2): p. 239-255.

Appendix 1

13.5dpc

17.5dpc

15.5dpc

Images of dissected mouse foetuses

Appendix 2

Foetal development

There are 26 stages in development and each of these stages, have clearly defined criteria which enables the researcher to identify the stage under investigation. When aiming for a range of stages, there are a number of factors to consider. Most importantly, the first morning a plug is noticed, is regarded as 0.5 dpc (0E), from this day to full term 17.5 - 19 days later, the embryo develops into a foetus at a rate determined by the mother's health and the number of pups in the uterine horns. The developing foetuses do not grow at a uniform rate and the resulting pups are often at slightly different developmental stages at a given time in the pregnancy. For example, if the uterine horns are dissected out at 13.5dpc, there will be a range of pups around this time within each horn. Clearly, when the pups are dissected out of their sacs it is easier to classify the developmental stage, however, if one is interested in the placenta and surrounding fluid then the embryo needs to be cryostated whilst in their sacs. Preliminary diagnosis of developmental stages comes prior to sectioning and requires the careful scrutinising of sections after cryostating, in addition to a surplus of sections to take into account those that are incorrectly diagnosed and cut at an odd angle.

The organs in mammalian development peak at different times during gestation and are often unrecognisable in the early stages, as the organ, which they later become. There follows a brief summary of the key characteristics, demonstrating the development of the main organs development, providing a reference for staging pups.

Differentiation of the **heart** begins with the splitting of the extra-embryonic mesoderm to form the intra-embryonic coelom at the 7 - 7.5 d.p.c. in the early headfold

stage embryo. Around 7.5 - 8 d.p.c. this division, forms the cardiogenic plate, spanning the ventral mid-line in the prospective pericardial region. However, it is not until 8.5 - 9 d.p.c that the primitive heart as an s shape can be seen to beat regularly with primitive red blood cells observed within the embryonic vasculature. During the period 9 - 10.5 d.p.c. there are little changes other than an accentuation of the s shape and the breakdown of a relatively small segment of dorsal mesocardium between the inflow and outflow regions of the heart, with the formation of the transverse pericardial sinus. At the 10 - 10.5 d.p.c. stage, the heart is the most prominent of all the organ and the first organ system to differentiate and function in the embryo.

The first major event in the differentiation of the **lungs** occurs at the 9 - 9.5 d.p.c. with the formation of the laryngo-tracheal grove, followed by the first indications of a pair of lung buds in the pericardio-peritoneal canals. At the 10 - 10.5 d.p.c. stage the canals extend either side of the heart, level with the lower part of the common atrial heart, in an uneven fashion initially with a single lobe on the left and four lobar bronchi on the right. By 12.5 - 13 d.p.c., the lungs have expanded to numerous segmental bronchi, consisting of homogenous cells, occupying a substantial volume of the peritoneal cavity. Detailed architecture of the future lung are first observed at 14.5 - 15 d.p.c., with the formation of terminal bronchioles dispersed throughout the lungs. During 15.5 - 17 d.p.c the lungs continue to branch such that by 17.5 - 18 d.p.c. the histological features resemble those at birth, with the exception of the cells that line the aveoli, which take on the squamous morphology at birth to facilitate gaseous exchange.

Primordial **germ** cells development is thought to originate during the early primitive streak stage (7 d.p.c.) as indicated by the presence of cells showing alkaline

phosphatase enzyme activity. In the slightly more advanced primitive streak stage these precursor germ cells are located in the mesodermal tissue of the viseral yolk sac, at the base of the allantois and at the caudal end of the primitive streak. By early somite stage (7.5 - 8 d.p.c) most primordial germ cells are located in association with the hindgut endoderm, some migrating to the mesentry of the hindgut by the 15 - 20 somite stage (9 - 9.5 d.p.c.), with the majority associated with the endoderm of the hindgut by somite stage 21 –30 (9.5 - 10.25 d.p.c.). By the 31 - 36 somite stage (10.25 - 10.5 d.p.c.) the majority of the primordial germ cells are found in the mesentry of the hindgut with a significant number having now migrated to the genital ridge.

Differentiation of the **urogenital** ridges, gonads and genital duct system begins around 9 d.p.c.. From 9.5 d.p.c. the urogenital ridges are first observed either side of the dorsal mesentry of the hindgut. By 11 d.p.c., the indifferent gonad can be seen, which increases in size and alkaline phosphatase activity over the following couple of days such that by 13 d.p.c. it is possible to distinquish between developing ovary and testis. At 13.5 d.p.c. early events associated with spermatogenosis occur including a number of type A spermatogonia in division, whilst the ovaries have a homogeneous cell morphology, first showing of egg clusters at 15.5 d.p.c.. By 17.5 - 18 d.p.c. the ovaries are in their final resting position with primary follicles and in the male the majority of the components of the reproductive system are recognisable.

Kidney differentiation is first noted with the appearance of mesonephric duct and tubules at 11.5 - 12 d.p.c., this system is later incorporated into the male reproductive system and completely regresses in the female, to be replaced by the formation of the metanephros by about 12.5 - 13 d.p.c. as an outgrowth of the ureteric bud. By 14.5 d.p.c.

primitive glomeruli are found throughout the early kidney and a well defined cortex and medulla region developed by 16.5 d.p.c.. By 17.5 d.p.c. the kidneys have ascended in relationship to the gonads (which have descended) and have a greater variety of cellular morphology.

The **extra embryonic** material of early stages (10 somites) consists of the amnion and the yolk sac. Not until 11.5 dpc is the placenta clearly seen, which matures over the following 6 - 7 days forming a labrynthine zone of trophoblast cells, endothelium cells and embryonic vessels.

In addition to the above criteria for organ development there is some organ movement within the confines of the developing foetus, to accommodate growing organs and tissues whilst they find their final resting place.

Appendix 1

13.5dpc

17.5dpc

15.5dpc

Images of dissected mouse foetuses

Appendix 2

Foetal development

There are 26 stages in development and each of these stages, have clearly defined criteria which enables the researcher to identify the stage under investigation. When aiming for a range of stages, there are a number of factors to consider. Most importantly, the first morning a plug is noticed, is regarded as 0.5 dpc (0E), from this day to full term 17.5 - 19 days later, the embryo develops into a foetus at a rate determined by the mother's health and the number of pups in the uterine horns. The developing foetuses do not grow at a uniform rate and the resulting pups are often at slightly different developmental stages at a given time in the pregnancy. For example, if the uterine horns are dissected out at 13.5dpc, there will be a range of pups around this time within each horn. Clearly, when the pups are dissected out of their sacs it is easier to classify the developmental stage, however, if one is interested in the placenta and surrounding fluid then the embryo needs to be cryostated whilst in their sacs. Preliminary diagnosis of developmental stages comes prior to sectioning and requires the careful scrutinising of sections after cryostating, in addition to a surplus of sections to take into account those that are incorrectly diagnosed and cut at an odd angle.

The organs in mammalian development peak at different times during gestation and are often unrecognisable in the early stages, as the organ, which they later become. There follows a brief summary of the key characteristics, demonstrating the development of the main organs development, providing a reference for staging pups.

Differentiation of the **heart** begins with the splitting of the extra-embryonic mesoderm to form the intra-embryonic coelom at the 7 - 7.5 d.p.c. in the early headfold

stage embryo. Around 7.5 - 8 d.p.c. this division, forms the cardiogenic plate, spanning the ventral mid-line in the prospective pericardial region. However, it is not until 8.5 - 9 d.p.c that the primitive heart as an s shape can be seen to beat regularly with primitive red blood cells observed within the embryonic vasculature. During the period 9 - 10.5 d.p.c. there are little changes other than an accentuation of the s shape and the breakdown of a relatively small segment of dorsal mesocardium between the inflow and outflow regions of the heart, with the formation of the transverse pericardial sinus. At the 10 - 10.5 d.p.c. stage, the heart is the most prominent of all the organ and the first organ system to differentiate and function in the embryo.

The first major event in the differentiation of the **lungs** occurs at the 9 - 9.5 d.p.c. with the formation of the laryngo-tracheal grove, followed by the first indications of a pair of lung buds in the pericardio-peritoneal canals. At the 10 - 10.5 d.p.c. stage the canals extend either side of the heart, level with the lower part of the common atrial heart, in an uneven fashion initially with a single lobe on the left and four lobar bronchi on the right. By 12.5 - 13 d.p.c., the lungs have expanded to numerous segmental bronchi, consisting of homogenous cells, occupying a substantial volume of the peritoneal cavity. Detailed architecture of the future lung are first observed at 14.5 - 15 d.p.c., with the formation of terminal bronchioles dispersed throughout the lungs. During 15.5 - 17 d.p.c the lungs continue to branch such that by 17.5 - 18 d.p.c. the histological features resemble those at birth, with the exception of the cells that line the aveoli, which take on the squamous morphology at birth to facilitate gaseous exchange.

Primordial **germ** cells development is thought to originate during the early primitive streak stage (7 d.p.c.) as indicated by the presence of cells showing alkaline

phosphatase enzyme activity. In the slightly more advanced primitive streak stage these precursor germ cells are located in the mesodermal tissue of the viseral yolk sac, at the base of the allantois and at the caudal end of the primitive streak. By early somite stage (7.5 - 8 d.p.c) most primordial germ cells are located in association with the hindgut endoderm, some migrating to the mesentry of the hindgut by the 15 - 20 somite stage (9 - 9.5 d.p.c.), with the majority associated with the endoderm of the hindgut by somite stage 21 –30 (9.5 - 10.25 d.p.c.). By the 31 - 36 somite stage (10.25 - 10.5 d.p.c.) the majority of the primordial germ cells are found in the mesentry of the hindgut with a significant number having now migrated to the genital ridge.

Differentiation of the **urogenital** ridges, gonads and genital duct system begins around 9 d.p.c.. From 9.5 d.p.c. the urogenital ridges are first observed either side of the dorsal mesentry of the hindgut. By 11 d.p.c., the indifferent gonad can be seen, which increases in size and alkaline phosphatase activity over the following couple of days such that by 13 d.p.c. it is possible to distinquish between developing ovary and testis. At 13.5 d.p.c. early events associated with spermatogenosis occur including a number of type A spermatogonia in division, whilst the ovaries have a homogeneous cell morphology, first showing of egg clusters at 15.5 d.p.c.. By 17.5 - 18 d.p.c. the ovaries are in their final resting position with primary follicles and in the male the majority of the components of the reproductive system are recognisable.

Kidney differentiation is first noted with the appearance of mesonephric duct and tubules at 11.5 - 12 d.p.c., this system is later incorporated into the male reproductive system and completely regresses in the female, to be replaced by the formation of the metanephros by about 12.5 - 13 d.p.c. as an outgrowth of the ureteric bud. By 14.5 d.p.c.

primitive glomeruli are found throughout the early kidney and a well defined cortex and medulla region developed by 16.5 d.p.c.. By 17.5 d.p.c. the kidneys have ascended in relationship to the gonads (which have descended) and have a greater variety of cellular morphology.

The **extra embryonic** material of early stages (10 somites) consists of the amnion and the yolk sac. Not until 11.5 dpc is the placenta clearly seen, which matures over the following 6 - 7 days forming a labrynthine zone of trophoblast cells, endothelium cells and embryonic vessels.

In addition to the above criteria for organ development there is some organ movement within the confines of the developing foetus, to accommodate growing organs and tissues whilst they find their final resting place.