THE GENE EXPRESSION UNDERLYING

EXPERIENCE DEPENDENT PLASTICITY IN THE

MOUSE BARREL CORTEX

SUBMITTED FOR THE DEGREE OF Ph.D.

BY

RICHARD ABRAHAM (B.Sc.)

JANUARY 2005

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

ACKNOWLEDGMENTS

I am extremely grateful to my supervisors – Prof. Kevin Fox and Dr. Pete Kille – for their continued help, enthusiasm and encouragement. Financial assistance was provided by the Biotechnology and Biological Sciences Research Council (BBSRC) without which this work wouldn't have been carried out – Thank You.

Cheers to everyone in Kevin and Pete's Lab whose patience I have tested over the years. In no particular order Huw, Geth, Vicki, Jen, Anna, Suresh, Steve Turner, Graham, Chrissie, Twinkle, Phil, Nick and Ele – your help was very much appreciated. Also thanks to Nick Bray for your advice.

Finally, I would like to dedicate this thesis to my family. Your love and support has given me the confidence to complete this piece of work. To Mam and Dad words can't express my thanks for everything you do for me – not just for this thesis but throughout my life! To my beautiful Louise you were the inspiration for this. Despite all you were going through at the time you still had the strength to encourage me. I know that together we can overcome anything. For our little girl Elodie – your smiling face was a constant reminder of what is truly important in life!!

ABSTRACT

Throughout the life of an organism the synaptic connections that link neurons are in a constant state of flux, being strengthened and weakened in response to external stimuli. This plasticity provides the central nervous system with a means of adaptability, allowing it to respond to changes in the external environment and is believed to underlie the processes of learning and memory (Lynch, 2004).

Neuronal plasticity can be studied using the paradigm of vibrissal deprivation in rodents (Fox, 2002). The mystacial vibrissae are represented in layer IV of the somatosensory cortex by clusters of neurons, termed barrels, in which the majority of neurons are responsive to stimulation of only the corresponding vibrissae (Woolsey and Van der Loos, 1970). The deprivation of the vibrissae results in a shift in responsiveness of neurons in deprived barrels to stimulation of surrounding spared vibrissae, and potentiation of responses of neurons in spared barrels to stimulation of the corresponding (principal) vibrissae ie experience dependent plasticity (Fox, 1992, Simons and Land, 1987).

It had previously been shown that, like other plasticity paradigms, gene transcription, mediated by the cyclic AMP Response Element Binding Protein (CREB), is associated with experience-dependent plasticity in the mouse barrel cortex (Barth et al., 2000). The finding that upregulation of a reporter gene, under the transcriptional control of CREB, following vibrissal deprivation conditions that induce plasticity, has been complementated with studies showing that deletion of CREB impairs plasticity in the barrel cortex (Glazewski et al., 1999).

Using microarray technology, this study analysed the changes in expression of endogenous genes to vibrissal deprivation patterns that induce neuronal plasticity. It was found that significant upregulation of a sub-set of plasticity related genes (PRGs) occurs, as in CRE reporter gene studies, early in the time course of deprivation that leads to plasticity. Real-Time PCR has confirmed the upregulation of a number of these genes.

Functional analysis has identified a significant over-representation of genes involved in protein synthesis from the group of PRGs. Further to this three of the PRGs have been identified as candidate plasticity genes in a number of paradigms. The products of these genes are key components of the extracellular matrix. This suggests that modifications to this fundamental part of the central nervous system may be the mechanism that results in experience dependent plasticity in the mouse barrel cortex.

CONTENTS

CHAPTER 2

MATERIALS AND METHODS

CHAPTER 3

USE OF MICROARRAY TECHNOLOGY: PROTOCOLS AND ANALYSIS

CHAPTER FIVE

HIGH DENSITY MICROARRAY ANALYSIS OF GENE EXPRESSION IN THE BARREL CORTEX FOLLOWING VIBRISSAL DEPRIVATION

CHAPTER SIX

FUNCTIONAL ANNOTATION OF GENES DIFFERENTIALLY EXPRESSED IN RESPONSE TO VIBRISSAL DEPRIVATION

CHAPTER SEVEN

GENERAL DISCUSSION

ABBREVIATIONS AND CONVENTIONS

The following abbreviations are used throughout the text.

$A_{(x)}$	Absorbance at x nm
ALL	Undeprived Mystacial Vibrissae
AP5	$_D$-2-amino-5-phosphonopentanoate
ATP	Adenosine Tri-Phosphate
bp	Nucleotide base pairs
BLAST	Basic local Alignment search Tool
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
C_t	Threshold Crossing
CaM	Ca^{2+} bound Calmodulin
CB	Chessboard Deprived Mystacial Vibrissae
CPG	Candidate Plasticity Gene
CRE	cyclic AMP Response Element
CREB	cyclic AMP Response Element Binding Protein
Cy	Cyanine
cAMP	Cyclic-3',5'-Adenosine Monophosphate
DAVID	Database for Annotation,Visualisation Integrated Discovery
dATP	Deoxy-Adenosine-Tri-Phosphate
dCTP	Deoxy-Cytosine-Tri-Phosphate
dGTP	Deoxy-Guanidine-Tri-Phosphate
dTTP	Deoxy-Thymine-Tri-Phosphate

dUTP	Deoxy-Uracil-Tri-Phosphate
DEP	All Mystacial Vibrissae Deprived
ds	Double stranded
DTT	Dithiothreitol
DNA	Deoxyribonucleic acid
EASE	Expression Analysis Systematic Explorer
EDTA	EthyleneDiamineTetracetic Acid
FAM	6-carboxyfluorescein
g	Gravity
gDNA	Genomic DNA
HEPES	(N-[2-HydroxyEthyl]-Piperazine-N'-[2-EthaneSulphonic acid]
IEG	Immediate Early Genes
kb	Nucleotide kilobase pairs
kDa	KiloDalton
LB	Luria Broth (Luria Bertani Media)
LTD	Long Term Depression
LTF	Long Term Facilitation
LTP	Long Term Potentiation
l-LTP	late phase Long Term Potentiation
M	Molar concentration
MAPK	Mitogen Activated Protein Kinase
MMLV-RT	Moloney-Murine Leukaemia Virus – Reverse Transcriptase
mRNA	Messenger RNA
N	Number of sample analysed
n	Nucleotide

NCBI	National Centre for Biotechnology Information
NMDA	N-Methyl-D-Aspartate
Nt	Nucleotide
OD	Optical Density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PKA	Protein Kinase A
PND	Post-natal day
PRG	Plasticity Related Gene
QPCR	Quantative Real –Time Polymerase chain reaction
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
ROX	Carboxy-X-rhodamine
RPM	Revolutions per minute
RT	Room temperature
RT-PCR	Reverse transcription-polymerase chain reaction
SAM	Significance Analysis of Microarray
SD	Standard Deviation
SDS	Sodium Dodecyl Sulphate
SEM	Standard error of the mean
Ss	Single stranded
SSC	Saline Sodium Citrate buffer
ssDNA	Single stranded DNA
STF	Short Term Facilitation
TAE	Tris Acetate-EDTA buffer

TAMRA	6-carboxy-tetramethyl-rhodamine
TE	Tris-EDTA buffer
T_m	Melting temperature
tRNA	Transfer RNA
U	enzyme Units
UTR	Untranslated region
UV	Ultraviolet
w/v	Weight by Volume
v/v	Volume by Volume

All laboratory materials employed are detailed according to their commercial name as given in the appropriate manufacturer's catalogue.

Genes are discriminated from proteins by the use of italicisation.

CHAPTER 1

GENERAL INTRODUCTION

1.1 ASPECTS OF NEURONAL PLASTICITY

The complex cellular architecture of the nervous system was largely unrecognised until the end of the 19th century. Prior to this time it was widely believed that the nervous system was a synctinium in which the cytoplasm of nerve cells was interconnected forming a continuous network. The reticular theory was dismissed mainly through the work of Ramon Y Cajal, a Spanish neurobiologist who proposed a network of individual cells connected at specialized sites, synapses (Cajal, 1952).

The brain is a dynamic organ capable of modifying in response to environmental input. This property - Neuronal Plasticity – can be studied at the molecular, through to the structural level in a variety of organisms. It is believed that, through mechanisms of neuronal plasticity, an organism learns from, and adapts to its environment. An elucidation of these mechanisms will provide insights into the working and function of the brain, arguably the most complex structure in the universe.

1.2 NEURONAL PLASTICITY IN THE RODENT BARREL CORTEX

The rodent barrel cortex is an area of the primary somatosensory cortex in which a clearly defined anatomical map of the sensory vibrissae can be observed. The mystacial vibrissae are represented in layer IV of the somatosensory cortex by clusters of neurons, termed barrels, in which the majority of neurons are responsive to stimulation of only the corresponding vibrissae (Woolsey and Van der Loos, 1970). Plasticity can be induced in the barrel cortex by the deprivation of the vibrissae which has two main effects; a shift in responsiveness of neurons in deprived barrels to stimulation of surrounding spared vibrissae, and potentiation of responses of neurons in spared barrels to stimulation of the corresponding (principal) vibrissae (Fox, 1992, Simons and Land, 1987). The unique topological arrangement of the barrel field means the locus of plasticity is clearly defined and vibrissal deprivation has become a paradigm for the study of experience dependent plasticity.

1.2.1 STRUCTURE AND FUNCTION OF THE RODENT MYSTACIAL VIBRISSAE

Rodents are nocturnal animals that rely on their mystacial vibrissae to perceive the external environment. The vibrissae are thin highly flexible hairs that sit in a follicle coupled to the skin, in an attachment that acts as a pivot, allowing the vibrissae to be moved (figure 1.1). Movement is controlled by a musculature framework (figure 1.1) that surrounds the follicle acting as a sling to propel the vibrissae forward. The subsequent retraction of the vibrissae is a passive process that is caused by the elasticity of connective tissue (figure 1.1) which supports the follicle at its base (Dorfl, 1982). The follicle of each vibrissa is innervated by 50-200 trigeminal sensory neurons (figures 1.1 and 1.2) which surround the follice providing free-nerve endings / lanceolate receptors and Merkel cell-neurite complexes (figure 1.2) which directly detect vibrissal movement (Renehan and Munger, 1986).

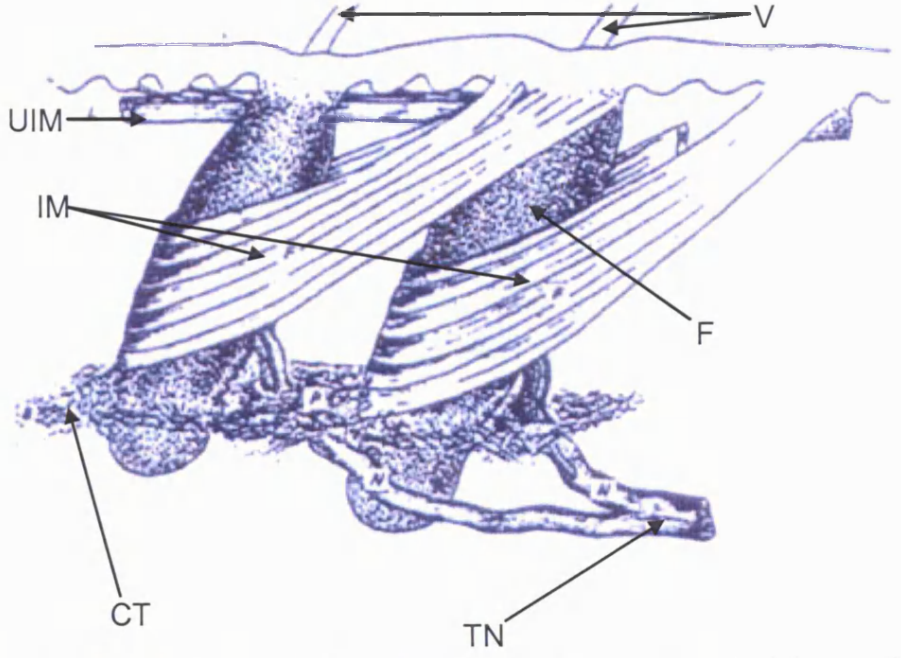


Figure 1.1 Musculature of the rodent vibrissal follicle. The intrinsic muscle (IM) forms a sling around the follicle (F) to pivot the vibrissae (V) forward. The trigeminal nerve, (TN) exits from the follicle. The upper intrinsic muscle (UIM) is anchored to the skin. A sheet of elastic connective tissue (CT) provides passive retraction in response to directed forward pivoting (adapted from Dorfl, 1982).

The degree to which rodents rely on their vibrissae in gathering information is evident from the highly ordered way the large mystacial vibrissae are arranged (figure 1.3). On the face of rodents is a muscular thickening, termed the mystacial pad, where the

vibrissae are located. Short vibrissae, a few mm long, are clustered around the mouth. The more posterior vibrissae are longer, about a few cm, and are arranged in a regular pattern (figure 1.3) (Waite, 1972). This stereotypical arrangement means that individual vibrissae can be identified from their respective position by arc letter and row number (figure 1.3) (Woolsey and Van der Loos, 1970). It is believed that the larger vibrissae provide spatial information about the immediate environment whereas the shorter vibrissae are used for discerning textures of objects (Brecht et al., 1997). The rodent uses a procedure known as whisking, in which movement of the vibrissae, across objects enables it to perceive the environment (Welker, 1964). Attempts have been made to quantify the level of discrimination afforded by whisking through texture discrimination testing, with the finding that the sensitivity is comparable to a primate fingertip (Carvell and Simons, 1990).

Whisking can be controlled by both the individual vibrissal muscles (figure 1.1) and the muscles that link the mystacial pad (Wineski, 1985). However, the general consensus is that vibrissae are whisked as a group and aberrant individual vibrissal whisking has not been shown (Vincent, 1912, Wineski, 1985). Whilst the protraction of the vibrissae is an active process directed by various muscles, the retraction depends on the elasticity of the connective tissue (figure 1.1) and it is this that limits the rate of whisking. Whisking is invoked by the trigeminal nerve (termed TN, figure 1.1) thought to be under the control of a brainstem pattern generator (Gao et al., 2001) which itself is subject to higher-level regulation from the motor cortex (Donoghue and Wise, 1982, Miyashita et al., 1994).

1.2.2 VIBRISSAE TO CORTEX SENSORY PATHWAY

The trigeminal sensory neurons that enter each vibrissal follicle detect movement of the vibrissae in response to whisking or when contact is made with an object. The cell bodies of these neurons are located in the trigeminal ganglion, where they form synapses with the infraorbital branch of the trigeminal nerve. Rapidly adapting action potentials can only be invoked in the sensory neurons through the deflection of a single vibrissa, no other vibrissa will elicit such a response (Gibson and Welker, 1983 a, b). Each sensory neuron responds in a particular way to vibrissal deflection. For example, some are relatively sensitive, requiring only minimum deflection to generate an action potential while others require a gross flickage. Regardless of the sensitivity of a neuron only a single (principal) vibrissa can cause this effect (Gibson and Welker, 1983 a, b).

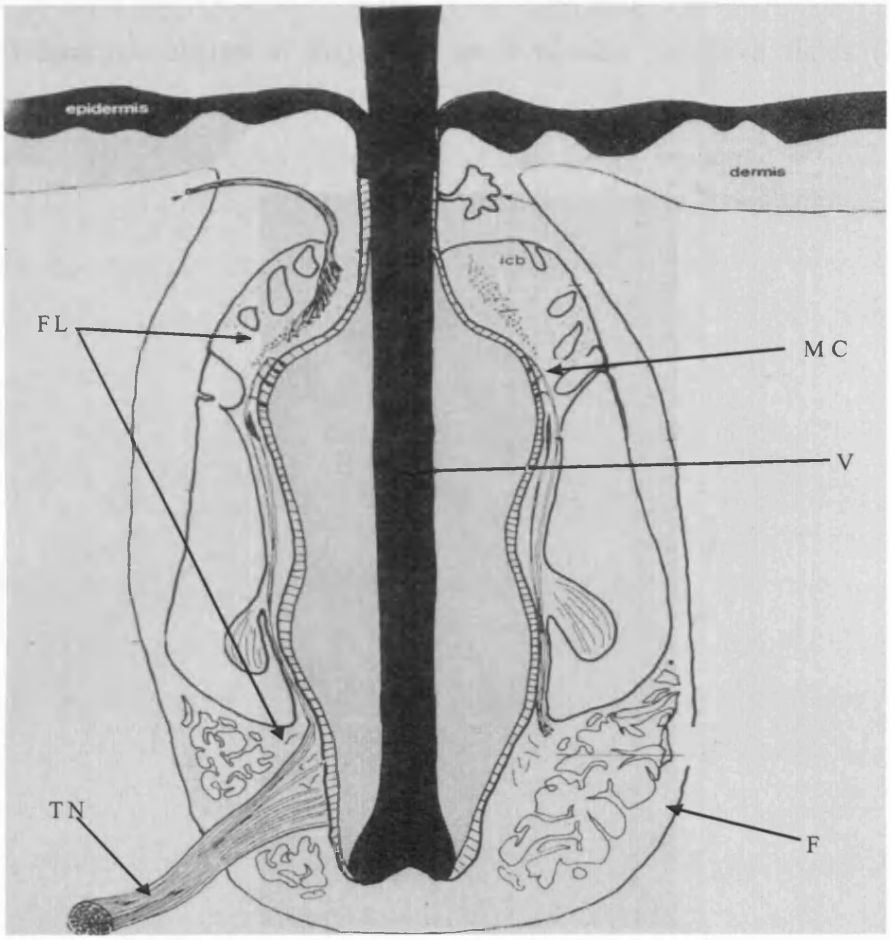


Figure 1.2 Sensory receptors of the mystacial vibrissae

The trigeminal nerve (TN) enters and divides, supplying free nerve endings / lanceolate receptors (FL) at the top and bottom of the follicle (F) and forms Merkel cell-neurite complexes (MC). Adapted from Renehan and Munger, 1986.

Signals are relayed from the infraorbital branch of the trigeminal nerve to the trigeminal brainstem nuclear complex where the nerve branches in two. One branch ascends and innervates the trigeminal nucleus principalis of the brainstem while the descending branch synapses with the three subnuclei of the spinal trigeminal nucleus - oralis, interpolaris and caudalis (figure 1.4, Agmon and Connors, 1991). The axonal projections from individual vibrissae form distinct clusters within trigeminal nucleus principalis which can be visualised. These clusters, or barrelettes are a somatotopic representation of the large mystacial vibrissae (Jacquin et al., 1993), made up of neurons that will only respond to their principal vibrissae, and have a single vibrissae receptive field (Ma, 1991, Veinante and Deschenes, 1999). This somatotopic representation is not evident in the

spinal trigeminal nuclei and likewise the neurons that project axons here are responsive to multi-vibrissal stimulation ie they have multi-whisker receptive fields (Jacquin and Rhoades, 1990).

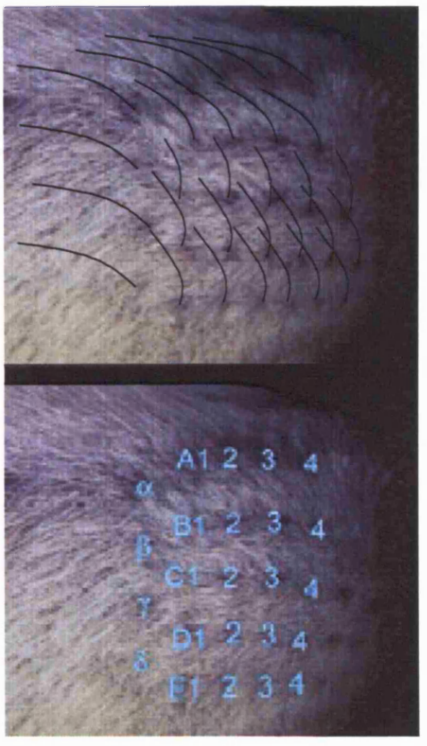


Figure 1.3 The Arrangement of the large mystacial vibrissae on the rodent snout.
The stereotypical arrangement of the large mystacial vibrissae and the nomenclature developed to identify individuals is provided. The vibrissae are arranged in arcs and rows. Five arcs are labelled a – e while a single arc containing the $\alpha - \delta$ vibrissae is generally regarded as a separate entity. For convenience only rows 1 – 4 are labelled in the figure, however rows C, D and E can contain up to 8 vibrissae. Adapted from Petersen, 2003.

The projections to the trigeminal nucleus principalis form the lemniscal sensory pathway and are thought to encode spatial information, while those to the spinal trigeminal nucleus comprise the paralemniscal pathway, proposed to relay temporal information (Ahissar et al., 2000). The pathways remain segregated with neurons from the trigeminal nucleus principalis projecting to the ventral posterior medial nucleus of the thalamus (figure 1.4). Again the segregation of the axonal fields is apparent with the somatotopic clusters or barreloids evident. While most neurons in a barreloid present single whisker receptive fields there are many multi-responsive neurons (Simons and Carvell, 1989, Nicolelis et al., 1993, Brecht and Sakmann, 2002a). Neurons from the spinal trigeminal nucleus project to the posterior medial nucleus and are broadly tuned with large multi-whisker

response fields (Diamond et al., 1992a). The destination of both ventral posterior medial nucleus and posterior medial nucleus neuronal projections is the primary somatosensory cortex (figure 1.4).

The ventral posterior medial nucleus afferents project primarily to layer IV of the neocortex with less dense terminations in layers 5/6. The afferents form clusters that can be visualised through staining and the somatotopic representation of the vibrissae is strikingly evident (see figure 1.5). In 1970 Woolsey and Van der Loos first used the term 'barrels' to define these structures and, following analysis of histological and electrophysiological data showed that a primarily one – to – one relationship exists between the large mystacial vibrissae and its corresponding barrel (Woolsey and Van der Loos, 1970, Woolsey et al., 1975). The area of layer IV containing the barrels is termed the barrel field or barrel cortex.

Neurons from the posterior medial nucleus do not form such clusters within the primary somatosensory cortex and afferents terminate in the layer IV septa (figue 1.4) between barrels and in layers II/III and V/VI. Reciprocal connections also exist (figure 1.4) from the barrel cortex to the primary motor cortex, and the secondary somatosensory cortex in addition to corticothalamic feedback circuits from the primary motor cortex to a number of thalamic nuclei (Porter and White, 1983).

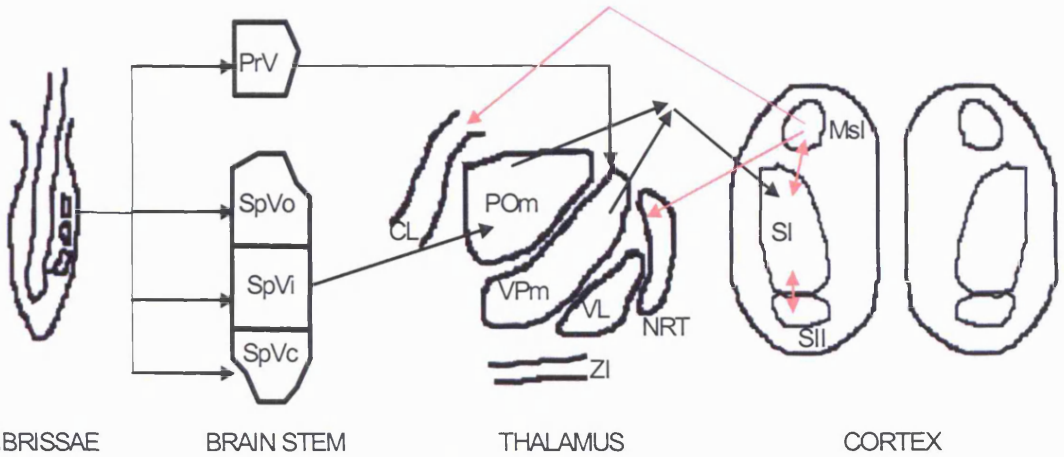


Figure 1.4 The neuronal projections from the vibrissa to the primary somatosensory cortex.

Receptive cells in the follicle of vibrissae project primary afferents to both the spinal trigeminal nuclei (SpVo, SpVi and SpVc) and trigeminal nucleus principalis (PrV) in the brain stem. SpVo, SpVi and SpVc afferents synapse at the posterior medial nucleus (POm) in the thalamus while PrV project to the ventral posterior medial nucleus (VPm) of the thalamus. Neurons from both the VPm and POm project to the primary somatosensory cortex (SI), VPm afferents terminate in discrete clusters in layer IV while those from the POm are dispersed between VPm clusters and in layers II/III and V/VI. Feedback circuits to the the thalamus from primary motor cortex (MsI) and intracortical reciprocal relays are shown in red.

1.2.3 INTRA-CORTICAL NEURONAL CONNECTIONS

In layer IV the primary excitatory input is from the ventral posterior medial nuclei neurons by glutamatergic synapses (Agmon and Connors, 1991). Neurons of the same barrel within layer IV are intimately linked, with connections every other pair of neurons, while between neighboring barrels there is very little connectivity (Feldmeyer et al., 1999, Petersen and Sakmann, 2000, 2001). This strengthens the ideas proposed by Woolsey and Van der Loos, that layer IV barrels are composed predominantly of neurons dedicated to the task of receiving information from their principal vibrissae.

The columnar idea of a barrel was proposed through the realisation that excitatory projections from layer IV neurons terminate predominantly in layers II/III of the same barrel (Feldmeyer et al., 2002). No reciprocal excitatory connections from layer II/III to layer IV have been observed, however layer II/III neurons show extensive coupling with neighbouring layer II/ III neurons and layer V/VI (Reyes and Sakmann, 1999). Thus the spread of neuronal excitation throughout the barrel field due to the deflection of vibrissae is probably mediated by neurons within layer II/III that are stimulated by layer IV neurons, localised to independent, tight-knit barrels that are intimately linked to a single vibrissae.

Prior to conducting vibrissal deprivation experiments Armstrong-James and Fox (1987) deemed it appropriate to accurately quantify the response properties of cells in all layers of the somatosensory cortex to principal and neighbouring vibrissal deflection. This was undertaken in order to clarify findings that seemed to erode the one-to-one vibrissae to barrel hypothesis. It had been suggested that neurons in layer IV, in addition to neurons in the supragranular and infragranular layers, received input from inappropriate vibrissae ie not the principal (Simons 1978, Welker, 1976). If this was the case then the degree of this inappropriate input had to be quantified otherwise changes in the receptive field properties of neurons, in response to vibrissal deprivation could have been misrepresented. Using a standardised deflection and a quantitative estimation of the receptive field size, based on the magnitude of response rather than subjective estimates, it was shown that 83% of layer IV cell stimulation is caused by the principal vibrissae (Armstrong-James and Fox, 1987). This number was similar to that found by Simons (1978).

As the vibrissae on the mystacial pad tend to work as a single unit (Vincent, 1912, Wineski, 1983), the barrel cortex is an ideal place for the integration and correlation of information from the individual vibrissae. The spatial and temporal distribution of this

information during cortical processing is being elucidated through the use of two techniques. In the first, extracellular electrodes are placed throughout the barrel field in all layers and responses to vibrissal deflection are recorded (Petersen and Diamond, 2000). The second technique uses voltage sensitive dyes in concert with high resolution imaging to record the profile of excitation flow (Petersen et al., 2003). The results of these techniques correlate well with traditional methods that measure the receptive field properties of many neurons when vibrissae are sequentially perturbed (Simons 1978, Armstrong-James et al., 1992, Zhu and Connors, 1999).

To date the temporal and spatial flow of excitation in a barrel is proposed to occur as follows; the first response to vibrissal deflection occurs at ~8ms in layer IV of the precise barrel corresponding to the deflected vibrissae and a few neurons in the mid-layer V/VI, followed a few seconds later by transmission within a column to layers II/III. Through local cortical interactions, excitation spreads within 12 milliseconds to layers II/III in neighbouring barrels and layer VI in the aligned barrel. Preference is given to barrels within a row such that deflection of D2 vibrissae, for example, will result in greater excitation in barrels D1 and D3, than either C2 or E2. After approximately 50ms excitation is distributed across the entire barrel field though the strength of response is significantly diminished (Petersen, 2003).

1.2.4 ARRANGEMENT AND CELLULAR COMPOSITION OF BARRELS

The striking similarity between the arrangement of barrels in layer IV and the large mystacial vibrissae can be seen by comparing figures 1.3 and 1.5. In figure 1.5 a cytochrome oxidase staining technique has been used to visualise the distinctive somatotopic map formed by thalamic afferents in layer IV. The well-defined barrel field is clearly visible and the relatively large area of cortex devoted to it highlights the dependency some rodent species have for their vibrissae (Welker, 1976, Woolsey et al., 1975). As rodents are nocturnal, visual input, as a means of perceiving the surroundings, will be limited, and greater reliance would be placed on the vibrissae. This is possibly the reason for the differences in both size and definition of the area of the somatosensory cortex devoted to processing vibrissal compared to visual information (figure 1.5)

The barrels representing the large mystacial vibrissae make up the posteriomedial barrel field (PMBSF). The stereotypical arrangement of the barrels in the PMBSF allows the nomenclature used to describe the vibrissae to be applied, such that each individual barrel can be identified by its arc (a – e) and row (1 ~ 9, dependent on the arc).

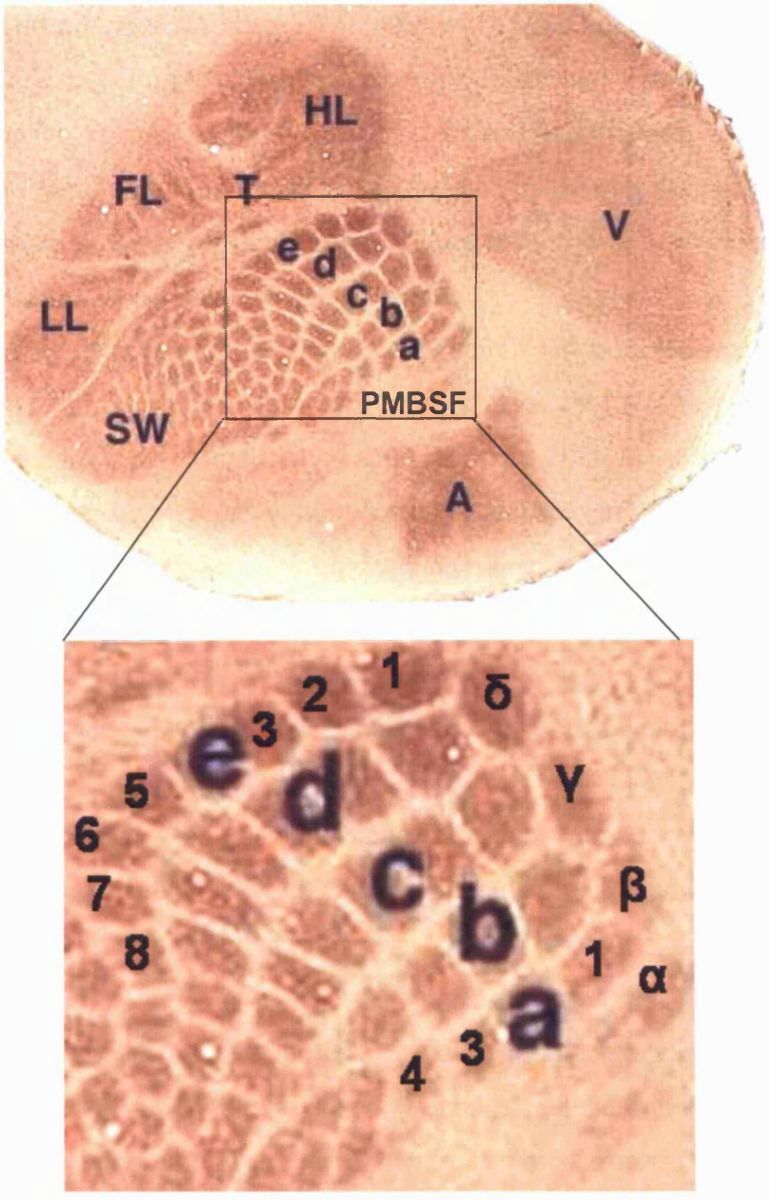


Figure 1.5 Cytochrome Oxidase stain of layer IV of the rodent somatosensory cortex. The barrels corresponding to the large mystacial vibrissae make up the posteriomedial barrel sub-field (PMBSF), enclosed by box. The smaller barrels corresponding to the short anterior vibrissae around the mouth (SW). Note the clearly defined cytoarchitectonic arrangement of barrels and the relatively large area of cortex devoted to them, compared to the visual (V) and auditory cortex (A). LL, FL and HL correspond to lower lip, forelimb and hindlimb respectively. This is an indicator of the rodents dependency on the vibrissae as a sensory tactile organ.

Inset – Expanded view of the PMBSF. Here the barrels of the large mystacial vibrissae are arranged in the same stereotypical pattern as the vibrissae. Therefore the naming covention applied to the vibrissae is used to define individual barrels. The barrels are arranged in arcs and rows. Five arcs are labelled a – e while a single arc containing the α – δ barrels is generally regarded as a separate entity. For covenience only row numbers for arcs a –e are labelled in the figure however the nomenclature is applicable to all. Adapted from http://www.neurobio.pitt.edu/barrels/pics.htm

The term barrel provides an accurate impression of the three-dimensional structure of these clusters. Using Golgi-staining techniques Woolsey and Van der Loos studied barrel structure in tangential and coronal sections of the somatosensory cortex. The larger barrels of the PMBSF are elliptical in shape and composed of a ring of high cell density, termed the barrel side, surrounding a less cellular area, the hollow. Distinct barrels are separated by an almost acellular area, the septum.

In the coronal plane the barrel sides are clearly delineated from the neighbouring barrels by the septa through the entire 100μm thickness of layer IV (Woolsey and Van der Loos, 1970). Larger barrels are composed of approximately 2000 neurons, of which 80% are located in the barrel side. The boundary between layers IV and III is less well demarcated compared to layers IV and the cell sparse layer V. Layer IV cells are composed predominantly of three classes of neuron (Simons and Woolsey, 1984);

pyramidal cells – with conical somata and stout apical dendrites

Class 1 neurons – Spiny stellate cells, with four to six dendrites, radiating from a round soma and cells with superficially directed dendritic process with the appearance of an apical dendrite.

Class II neurons - One type of cell has a bipolar appearance with cylindrical dendritic field projecting to layers V and VI in addition to the supragranular layers, the other has upwards of 10 dendrites which arborise throughout layer IV.

1.2.5 DEVELOPMENTAL PLASTICITY OF THE BARREL CORTEX

Rodents are born with an immature central nervous system. The migration of cortical neurons is incomplete and newborn mice fail to show any evidence of barrel formation (Miller, 1988, Van der Loos and Woolsey, 1973). During early periods of life there is a critical period in the development of the somatosensory cortex, when altering the sensory input from the periphery has a profound effect on the anatomical arrangement of somatotopic maps. This developmental plasticity can be induced by cauterising the mystacial vibrissae which has profound effects on the shaping of the barrel cortex (figure 1.6) (Woolsey and Wann, 1976).

In a series of experiments the C-arc of one set of mystacial vibrissae were cauterized at post-natal days (PND) 1, 2, 3, 4, 5, 7, 10 and 20 while the other set were untreated.

Following sacrifice at PND 60 both barrel fields from individual animals were processed to reveal the barrel field in layer IV of the somatosensory cortex. In undeprived controls normally developed barrel fields with the stereotypical topography of rows and arcs that represent the vibrissae were observed. However, the barrel fields of the hemispheres whose contralateral C-arc vibrissae had been deprived at PND 1, 2, 3, 4 and 5 showed gross defects in this area of the barrel cortex (figure 1.6) (Woolsey and Wann, 1976). Cauterising at PND 1 results in complete ablation of the barrels whose vibrissae have been deprived with the space left occupied by barrels of neighboring arcs. If deprivation is carried out at PND 3-5 rudimentary arcs are observed, however the characteristic septa that separates barrels is not. Deprivation at PND 7 onwards had no observable impact with the barrel fields indistinguishable from untreated controls. These observations confirmed the ideas of Hubel and Wiesel (Hubel and Wiesel, 1970) that synaptic connections are highly malleable during critical periods of animals life. Alterations to sensory inputs on the periphery can result in changes to the somatotopic outlay of the sensory cortex that last for the lifetime of an animal.

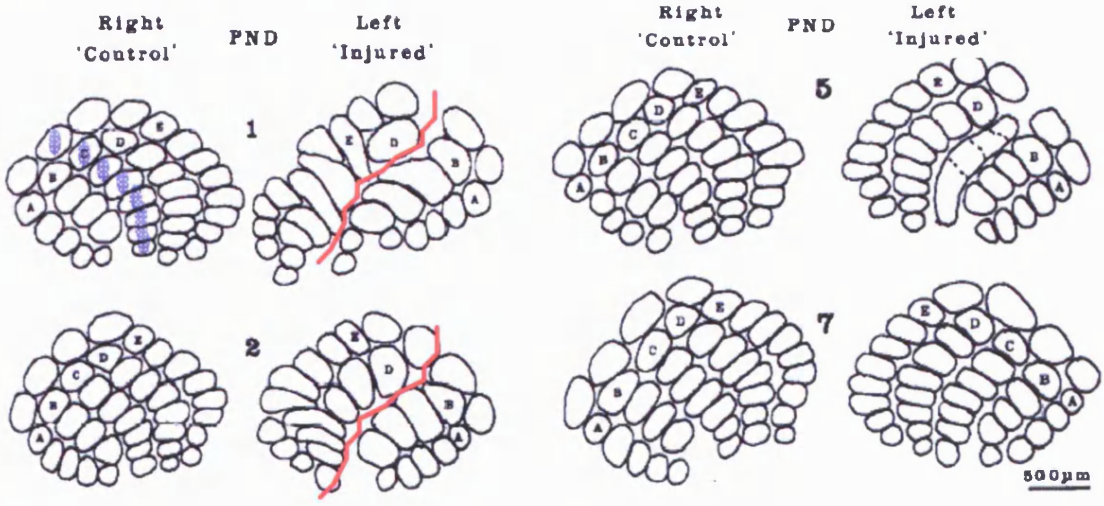


Figure 1.6 The effect of vibrissal deprivation on the development of the barrel cortex in newborn rodents. On one set of the large mystacial vibrissae the C-arc was deprived at post-natal days (PND) 1, 2 5 and 7, while the remaining set were untreated. At PND 60 the animals were sacrificed and tangential sections of the cortex were suitably prepared to view the barrel field in layer IV. Right undeprived 'control' animals had normally developed barrel fields showing the stereotypical arcs and rows relating to the pattern of the vibrissae (C-arc is highlighted blue in a typical control, left hand side figures). Animals where the C-arc had been deprived at PND 1, 2 and 5 showed significant disruption to the layout of the contralateral barrel field compared to controls, including complete abaltion of barrels at PND1 and PND2 (red). At PND7 vibrissal deprivation had no effect and the barrel field resembled that of a control. Adapted from Woolsey and Wann, 1976

These results revealed an interesting temporal correlation between the critical period for plasticity and the normal development of the PMBSF, where the cortical plate doesn't differentiate into layers until PND 6 (Weller 1972, Rice and Van der Loos, 1977). It has further been shown that thalamic afferents cluster to form barrels in layer IV during the first 4-5 post-natal days and barrel hollow and wall formation follows a similar time course (Erzurumulu and Jhaveri, 1990, Rice et al., 1985). If the vibrissae are plucked, as opposed to cauterized no developmental plasticity is observed (Weller and Johnston, 1975). Given that cauterisation disrupts the entire follicle including sensory cells of the trigeminal neuron, intact neuronal processes from the vibrissae to the trigeminal ganglion, with or without corresponding activity, are required for normal cortical development.

1.2.6 EXPERIENCE DEPENDENT PLASTICTY OF THE BARREL CORTEX

The gross effects of vibrissal cauterisation observed by Woolsey and Wann, (1976) are in contrast to the more subtle plastic changes that occur in the mouse barrel cortex following vibrissal trimming or plucking. Here the follicle of the vibrissae is left intact by the experimental procedure removing the ambiguity of whether the observed cortical plasticity is a response to the destruction of the periphery or a result of sensory experience (Fox, 1992).

The plasticity observed following vibrissal plucking, or trimming, is regarded as a change in the cortical domain over which the stimulation of vibrissae evoke responses. This phenomenon was intitially investigated by Simons and Land, (1987) in experiments that deprived entire rows of vibrissae, by trimming, from the day of birth. When the animals reached maturity (PND 45-60) the vibrissae were regrown and reponses to both principal and surround vibrissae were recorded for deprived and undeprived barrels in layer IV. Neurons in barrels that are undeprived respond with characteristic vigour to their principal vibrissae and poorly to adjacent vibrissae. Neurons in barrels that had had their principal vibrissae deprived showed significantly higher responses to adjacent vibrissal deflection. Changes are observed in the receptive field size, angular tuning and responsiveness of neurons in deprived barrels that are irreversible months after re-growth (Simons and Land, 1987). Given that the cortex is not completely developed until six weeks after birth the results provided evidence that plasticity can occur on a more subtle basis through procedures as innocuous as vibrissal trimming or plucking. Shaping of the sensory cortex through the hardwiring of synaptic connections occurs in direct response

to sensory experience. Vibrissal deprivation is one way of manipulating sensory input to induce experience dependent plasticity and has become a paradigm for its study.

As vibrissal cauterising does not induce developmental plasticity of the barrel cortex after PND 7 (Woolsey and Wann, 1976), Fox (1992) conducted experiments to discover whether experience dependent plastic changes followed a similar time course. In a variation on Simons and Land's paradigm a univibrissal approach was used in which all whiskers in experimental animals were deprived apart from the D1 vibrissa.

Deprivations were conducted starting at PND 0, 2, 4 and 7 days and revealed a critical period for which layer IV neurons display plasticity. If deprivations are carried out at PND 0 and 2, cells in barrels surrounding D1 will show a depression in response to deflection of their principal vibrissae and a potentiation of response to the D1 spared barrel, the archetypal plastic response. After PND 4 vibrissal deprivation does not induce plasticity in layer IV neurons (figure 1.7, bottom panel).

The spread of responsiveness of layer IV cells in deprived barrels to deflection of the D1 whisker is shown in figure 1.8 for controls and D1 spared animals. In the top panel it can be seen that rapid responses to D1 deflection in controls are confined to the D1 barrel. This is contrast to the D1 spared animals, bottom panel, where rapidly responding cells are located in neighbouring deprived barrels in addition to the D1 barrel.

In layers II/III plasticity is observed regardless of the day at which deprivations are started (figure 1.7, top panel). In contrast to layer IV cells the continued ability of the layer II/III cells to undergo plastic changes in response to vibrissal deprivation, regardless of the PND at which it is carried out, has been confirmed in 2-deoxyglucose studies (Hand, 1982).

Thus, experience dependent plasticity has a critical period when it can be invoked in layer IV neurons of the barrel cortex. This critical period corresponds to a time in the animals' life when the barrel cortex is not yet defined and thalamocortical afferents are just starting to segregate into the clusters that define barrels (Erzurumulu and Jhaveri, 1990, Senft and Woolsey, 1991). Increased responses of neurons in deprived barrels to undeprived vibrissae is thought to be due to a lack of activity in the deprived vibrissa - thalamic afferents that prevent synaptic connections from forming. Comparisons can be made with visual cortex plasticity observed in the cat where initially overlapping genitocortical afferents segregate to respond to either the left or right eye during the critical period. Blocking activity in either eye results in a bias of afferents that connect with the unblocked eye (Hubel et al., 1977). In the barrel cortex thalamocortical afferents synapse

with layer IV neurons. Through co-ordinated firing the synapses made by afferents from each vibrissae would be strengthened. Under Hebb's Law (Hebb, 1949) if multiple afferents are active simultaneously, as they would be if they were driven by the same vibrissae, then synapses between layer IV cells and thalamic afferents for the principal vibrissae would become strengthened. If the afferent input is disrupted, through vibrissal deprivation, synapses between neurons (in all layers) of the same barrel would be weakened, while those with inappropriate spared barrels would be strengthened.

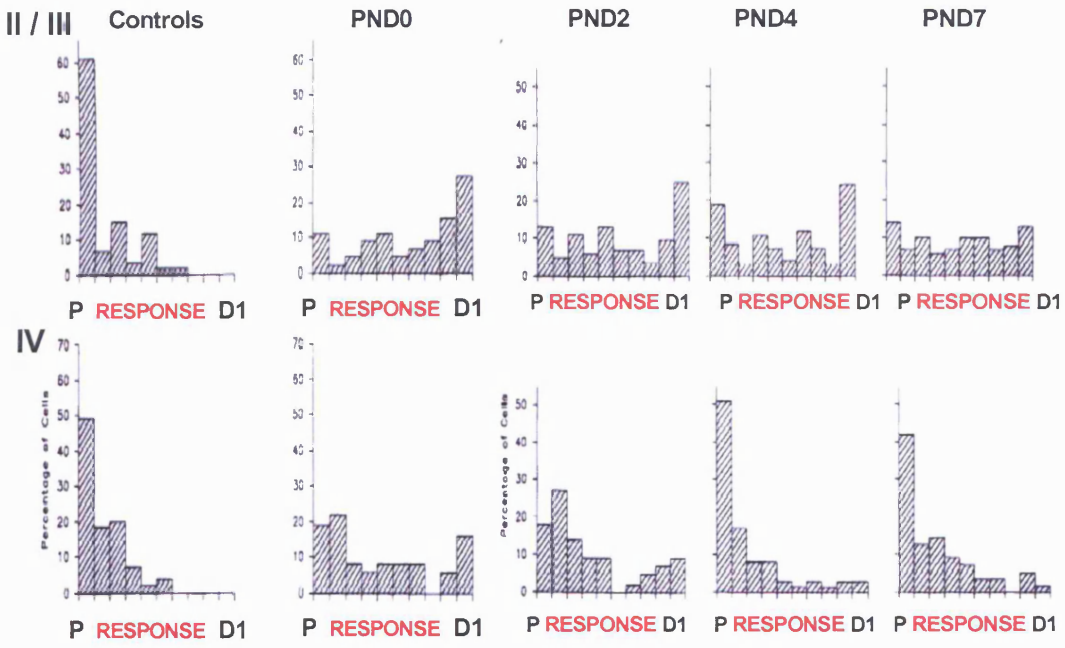


Figure 1.7 Evidence for a Critical Period in Experience Dependent Plasticity in the Rodent Barrel Cortex. Each chart plots the response of cells in layers IV (bottom) and II / III (top) to stimulation of the principal (P) and deprived (D1) vibrissae following univibrissal deprivation. In controls the majority of cells respond only to deflection of the principal vibrissae. Plasticity is induced in animals deprived at post-natal day (PND) 0 in layers II/III and IV, as indicated by a shift in reponsiveness of cells from the principal to the undeprived D1 vibrissae. In animals where deprivation was carried out later in life, post-natal day 4+, plasticity is no longer observed in layer IV and cells respond primarily to their principal vibrissae. In layers II/III plasticity is maintained even when deprivation is delayed. Adapted from Fox, 1992.

Puberty in rodents is generally considered to occur at about 1-2 months when major cortical development is complete. At this age plasticity is reduced in all layers compared to neonates, and layer IV cortical plasticity, both developmental and experience dependent, is lost completely (Fox, 1992, Glazewski and Fox, 1996). Layer II/III experience dependent plasticity however is still robust and the use of pubescent rodents in

plasticity experiments has the advantage that it can be studied without the confounding effects of the large-scale reorganisation of the barrel cortex that occurs in neonates (Fox, 1992). Furthermore plasticity is characterised by two processes; the potentiation of responses to spared vibrissae and a down-regulation (depression) of responses to re-grown deprived principal vibrissae (Glazewski et al., 1998).

In adults experience dependent plasticity has been observed in layers II/III until 15 months of age (Fox, 2002) though the depression of responses observed in adolescents ends after 2 months (Glazewski et al., 1998).

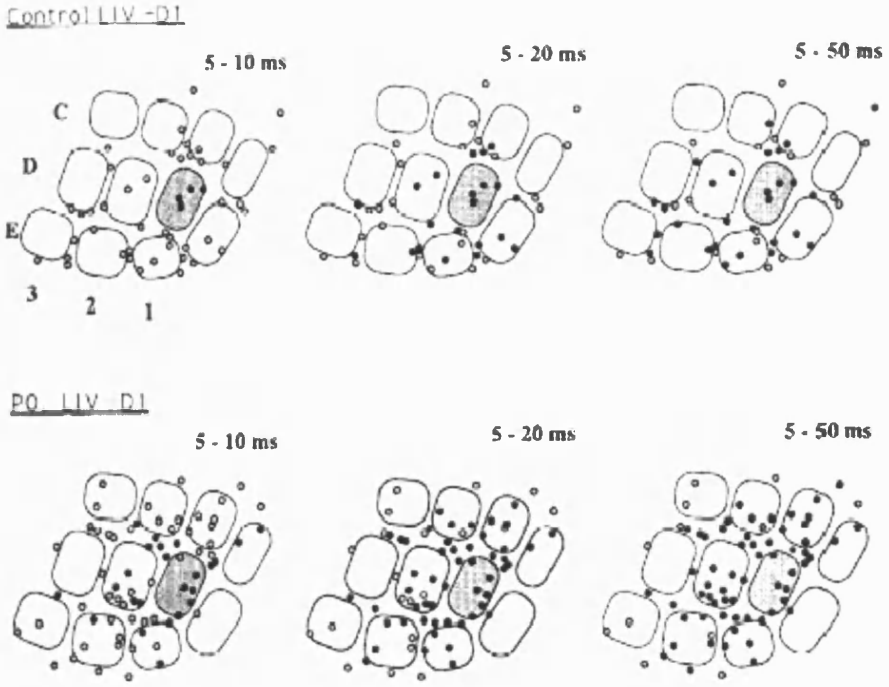


Figure 1.8 Effect of vibrissal deprivation on somatotopic representation in the mouse barrel cortex. This figure compares the cortical layer IV area over which a single vibrissal deflection is capable of eliciting a response. The D1 barrel is shaded. Solid circles represent the location of cells where a cell respond to deflection of the D1 vibrissae within 5-10ms, 5-20ms and 5-50ms (left middle right respectively). Empty circles are cells that elicited no response. The top figure is from recordings in undeprived controls, the bottom figure represents animals in which all vibrissae, except D1, were deprived at post-natal day 0. Adapted from Fox, 1992

Various manipulations of the large mystacial vibrissae have been used to further understanding of the physiological changes that underlie barrel cortex plasticity;

Single vibrissa deprived - Has been used to study depression effects as depression is enhanced in deprived barrels surrounded by undeprived barrels (Glazewski et al., 1998)

Single vibrissa spared - Has provided information that depression effects appear to depend on the proximity of the deprived and undeprived vibrissae (Glazewski and Fox, 1996)

Chessboard deprivation – Every other vibrissae is deprived and plasticity (both potentiation and depression of responses) is invoked across the entire barrel field (Wallace and Fox, 1999). The advantage of this deprivation is that pharmacological agents can be placed over the entire locus of plasticity, a potential problem if plasticity is only invoked in a single barrel.

1.3 THE MOLECULAR BASIS OF NEURONAL PLASTICITY

1.3.1 LONG TERM POTENTIATION

The sharing of ideas and knowledge amongst researchers has expanded understanding of the mechanisms of neuronal plasticity. This is perhaps best exemplified at the molecular level, where the universal nature of biomolecular interactions allows findings in one paradigm to be applied to another. Understanding of the cellular and molecular mechanisms that alter the efficiency of synaptic connections has been brought about largely through the study of long term potentiation (LTP) (Lomo, 1966). Studies have revealed that gene expression and *de novo* protein synthesis, directed by a number of signalling pathways (Stanton and Sarvey, 1984, Frey et al., 1996) are required for the maintenance of LTP.

Discovered in the 1960's (Lomo, 1966), LTP is a sustained increase in the efficiency of synaptic transmission which occurs in response to heightened neuronal activity. Thus, LTP, and its opposing process long term depression (LTD), are possible mechanisms whereby the central nervous system can alter its synaptic connectivity without the requirement for the generation of new neurons. LTP and LTD are therefore widely regarded as the cellular basis of learning and memory. LTP can be invoked in many regions of the brain, including the amygdala (Dityatev and Bolshakov, 2005), cerebellum (D'Angelo et al., 2004) and the cortex (Fox, 1995). LTP was first discovered in the

hippocampus, a structure known to be important for declarative memory (Scoville and Milner, 1957, Squire et al., 1984, Morris et al., 1990). The hippocampus is a curved elevation of gray matter which extends the entire length of the floor of the temporal horn, an area within the lateral ventricle of the neocortex. When sectioned transversely the neuronal circuitry is left intact and the lamellar structure makes it highly amenable to experimentation (Lomo, 1966).

LTP is principally studied on the synaptic connections between the Schaffer collaterals and CA1 neurons of the hippocampus (figure 1.9, top). In order to induce LTP a brief train of high-frequency stimulation is applied to the Schaffer collaterals. LTP is observed as an increase in the excitatory post-synaptic potentials (EPSPs) of the CA1 neurons which can persist for several hours when induced with low-frequency stimulation. This elevation is in direct contrast to the moderate, transient rise in EPSPs invoked by low-frequency tetanii without prior high frequency stimulation (figure 1.9, bottom) (Bliss and Lomo, 1973).

What makes LTP a good candidate for a mechanism by which memory is stored is that it fulfills the criteria set out in Hebb's Law (Hebb, 1949). The coincident – detection theory proposed by Hebb (1949) states;

repeated activation of a cell B by cell A, via a synaptic connection, elicits some type of metabolic change within one or other or both cells. These changes enable subsequent activations of cell B, by cell A, to occur more readily.

In simplistic terms this means *cells that fire together, wire together* or, the more a particular synapse is used, the stronger it becomes ie through persistent activation of the postsynaptic neuron by the presynaptic cell a sustained increase in the efficiency of the synapse is observed. Further, the synapse specific nature of LTP provides a highly selective method of fine tuning neural networks in response to the high demands life experience would place on it.

In vivo, if synaptic potentiation continued to increase unchecked a level of maximum efficacy would be reached that would make it difficult to encode new information. Long-term depression (LTD) has been proposed as a control mechanism that selectively weakens specific sets of synapses. Like LTP, LTD was initially identified at the synapses between the Schaffer collaterals and the CA1 pyramidal cells in the hippocampus. LTD occurs when the Schaffer collaterals are stimulated at a low rate for long periods (10–15 minutes). This pattern of activity depresses EPSPs for several hours (Barrionuevo et al., 1980, Fujii et al., 1991). LTP can erase the decrease in EPSP size due to LDP, and,

conversely, LDP can abolish the increase in EPSP size due to LTP. This suggests that both LTP and LDP can function at the same site providing a level of control that enables new information to be encoded.

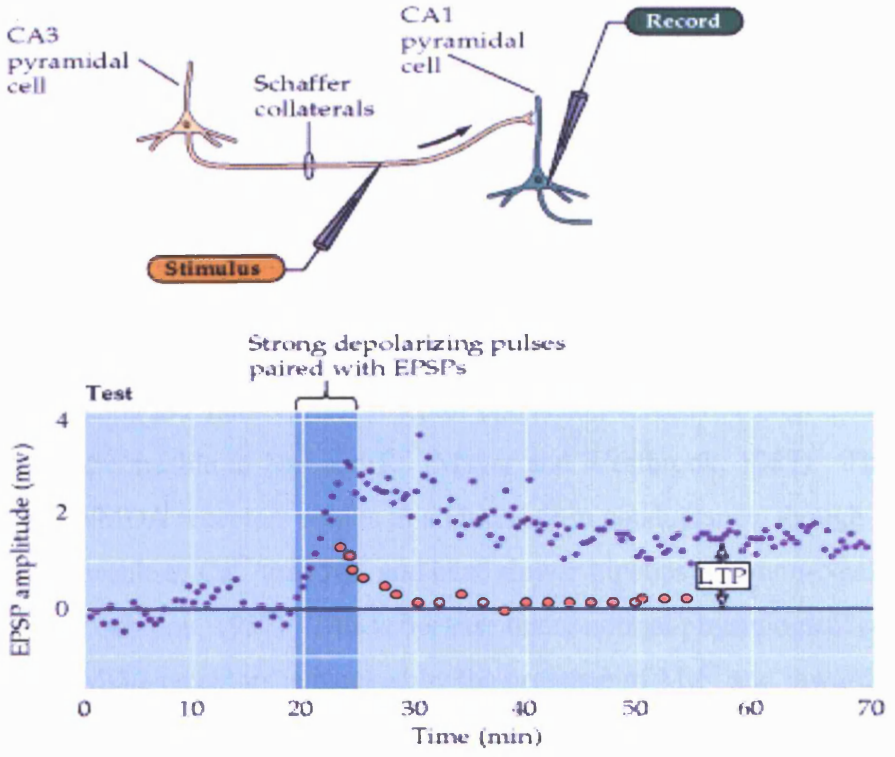


Figure 1.9 Induction of LTP in the CA1 region of the hippocampus

Top - Afferent excitatory Schaffer collaterals are subjected to brief trains of high-frequency stimulation. Recordings are taken, before and after stimulation, of the excitatory post-synaptic potential (EPSP) of CA1 neurons.

Bottom Size of the EPSPs are plotted before and after stimulation. When stimulation occurs every two to three minutes a brief rise in evoked EPSPs occurs before a return to baseline (hypothetical red circles). If a brief, high frequency train of stimuli is applied the size of the EPSPs is increased and potentiation lasts up to 20 hours (Gustafsson et al., 1987).

1.3.2 NMDA RECEPTORS AND NEURONAL PLASTICITY

Studies into the molecular aspects of neuronal plasticity essentially started in the mid 1980's when the structure and function of the NMDA receptor, an ionotropic glutamate receptor was elucidated. Glutamate is the primary neurotransmitter of the central nervous system that relays signals from the presynaptic neuron to the postsynaptic membrane across the synaptic cleft (Michaelis, 1998). At the postsynaptic membrane specialised

receptors bind glutamate, resulting in an exchange of ions across the post synaptic membrane. The resulting neurotransmitter-induced current flow alters the conductance and usually the membrane potential of the postsynaptic neuron, increasing or decreasing the probability that the neuron will fire an action potential (Purves et al., 2001).

The characteristics of the ionotropic glutamate receptors are determined by the composition of five sub-units from which they are formed. The group is subdivided into NMDA and non-NMDA receptors depending on their ability to bind the glutamate analogue N-Methyl-D-Aspartate (NMDA) (Verdoorn et al., 1991, Moyner et al., 1992; Nakanishi, 1992). The NMDA receptors are built from combinations of 5 subunits, NR1, NR2a – NR2d, which can co-assemble to form homomers or heteromers, though a functional NMDA receptor must have at least one NR1 subunit (Moyner et al., 1992, Nakanishi, 1992, Ishii et al., 1993).

Glutamate binding to a non-NMDA receptor opens non-selective cation channels that are preferentially permeable to Na^+ and K^+ (Mayer and Westbrook, 1987). The binding of glutamate to NMDA receptors results in an increase in conductance, though the receptors are more permeable to Ca^{2+} than Na^+ and have slower kinetics than non-NMDA receptors (Mayer and Westbrook, 1987). This is because under normal physiological conditions the ion pore in NMDA receptors is blocked by the presence of Mg^{2+} and inward current flow is prevented. Therefore low frequency synaptic transmission releases glutamate, which crosses the synaptic cleft and binds to both NMDA and non-NMDA receptors, though current will only flow through the non-NMDA channel. In order for NMDA receptors to function the Mg^{2+} block must be lifted and this occurs when the postsynaptic membrane is sufficiently depolarised (Nowak et al., 1984).

In LTP depolarisation of the postsynaptic membrane, through stimulation of the Schaffer collaterals, is sufficient to release the Mg^{2+} block from NMDA receptors. The resulting inward Ca^{2+} flow could account for the observed, sustained synaptic potentiation (Kelso et al., 1986, Wigstrom and Gustaffson, 1986).

The role of NMDA receptors in LTP was characterised predominantly through the use of pharmacological inhibitors which selectively affect receptor function. A number of antagonists are recognised; the glutamate analogue $_D$-2-amino-5-phosphonopentanoate (AP5) and channel blocker MK-801 act directly on the receptor. Their application to hippocampal slices prevents LTP (Collingridge et al., 1983, Coan et al., 1987).

In the rodent barrel cortex both developmental plasticity and experience dependent plasticity appear to be mediated through NMDA receptors. In newborn rats the normal

development of the cytoarchitecture of the barrel field is not impaired when NMDA receptor function is disrupted (Schlaggar et al., 1993). However if the vibrissae are cauterised early in life there is significant disruption to the somatotopic barrel field (Woolsey and Wann, 1976). In vibrissal deprived animals that are simultaneously implanted, with elvax releasing AP5 onto the surface of the cortex, the barrel field morphology develops normally. This suggests that NMDA receptors play a part in the developmental plasticity of the barrel field.

Further to this Rema et al., (1998) studied the effects of AP5 on experience dependent plasticity in rats of 2 – 3 months in age. The application of AP5 completely abolished the potentiation of neurons, in layer IV and layers II/III of deprived barrels, to spared vibrissae and also further reduced the depression of responses in deprived barrels (Rema et al., 1998). (Note: this experiment used a paired whisker paradigm to invoke plasticity. Hence, plasticity is observed in layer IV animals at later ages, compared with the univibrissal deprivation technique in which plasticity is not expressed in older animals (Fox, 1992)).

Both barrel cortex plasticity and LTP appear to be mediated through the NMDA receptor. Under Hebb's law unless both presynaptic and postsynaptic neurons are coincidentally potentiated, synapses are not strengthened (Hebb, 1949). NMDA receptors are only activated when both pre- and postsynaptic neurons are potentiated and can act in a ligand gated and voltage dependent manner. This has the potential for them to act as molecular co-incidence detectors, as they respond to presynaptic (glutamate) and postsynaptic (Mg^{2+} block) stimulation and is the reason that their involvement in neuronal plasticity has been studied so extensively. It appears that NMDA receptors are the conduits that signal to the post-synaptic neuron that both the pre and postsynaptic membrane are simultaneously depolarised. The nature of this signal is considered to be the influx of Ca^{2+} into the postsynaptic neuron triggered by the removal of the Mg^{2+} block.

1.3.3 GENE TRANSCRIPTION AND PROTEIN SYNTHESIS IN NEURONAL PLASTICITY

In the 1980's experiments were conducted to determine whether LTP was dependent on protein synthesis. This was in reaction to observations that increased secretions of newly synthesised proteins and increases in specific protein fractions accompanied LTP, though it was uncertain whether these observations were a by-product of or a direct requirement for LTP (Browning et al., 1979). It was found that LTP can effectively be partitioned into two components, an initial short phase that lasts approximately 3 hours that is unaffected

by protein synthesis inhibitors and a longer phase (l-LTP) that requires *de novo* protein synthesis (Stanton and Sarvey, 1984, Krug et al., 1984, Frey et al., 1988). It was further shown that administration of the RNA synthesis inhibitor actinomycin D also blocks l-LTP (Frey et al., 1996) suggesting that a transcriptional response underlies long-term synaptic potentiation.

The findings that gene transcription and de novo protein synthesis are required for LTP draw parallels with investigations in a plasticity paradigm that uses the marine mollusk *Aplysia californica*. Eric Kandel focused his interests in the molecular aspects of learning and memory in this invertebrate because of its simpler nervous system. Aplysia has only a few tens of thousands of neurons, many of which are quite large (up to 1 mm in diameter) and in stereotyped locations within the ganglia that make up the animal's nervous system making it practical to monitor the electrical and chemical signaling of specific, identifiable nerve cells (Kandel and Schwartz, 1982).

In *Aplysia*, a light touch to the animal's siphon results in gill withdrawal, a response that gradually habituates with repeated stimulation. After several repetitions, the animal no longer bothers to withdraw the gill after being touched. However, if touching the siphon is paired with a strong electrical stimulus to the animal's tail, then the siphon stimulus again elicits a strong withdrawal of the gill. The noxious stimulus to the tail sensitizes the gill withdrawal reflex to light touch. Even after a single stimulus to the tail, the gill withdrawal reflex remains enhanced for at least an hour (called short-term sensitization the equivalent of short-term LTP). With repeated pairing of tail and siphon stimuli, this behaviour can be altered for days or weeks. Such a long-lasting change in the gill withdrawal reflex is an example of long-term sensitization/facilitation (LTF) a paradigm that correlates with l-LTP in the hippocampus (Kandel and Schwartz, 1982). As observed in LTP, LTF is also dependent on gene transcription and *de novo* protein synthesis (Castellucci et al., 1986).

1.3.4 Ca^{2+} SIGNALLING PATHWAYS LEADING TO GENE TRANSCRIPTION

The earliest identified genes that could be linked to LTP were the immediate early genes (IEG) *zif268* and *c-fos* related genes (Richardson et al., 1992, Wisden et al., 1990). In general IEG are activated in a wide range of tissues under varying stimuli by a process that does not require *de novo* protein synthesis. Estimates put the number of IEG that are functional at the synapse to approximately 30 with roughly half being transcription factors (Lanahan and Worley, 1998). Upregulation of both the mRNA and functional

proteins of many IEG accompany short term and l-LTP. The two best studied IEG are *zif268* and *Arc* which have quite different properties.

Transgenic mice in which *zif268* has been deleted show impaired l-LTP though its lack of upregulation in paradigms that invoke intermediate LTP suggest its function is to consolidate and maintain potentiation (Hughes et al., 1999).

Arc (activity-regulated cytoskeleton-associated protein) mRNA and protein is upregulated by synaptic activity dependent on NMDA signalling and inhibition of LTP is caused by disruption of *Arc*. The inhibition of LTP may be related to the fact that *Arc* interacts with the cytoskeleton through actin suggesting a remodelling of neuronal infrastructure may underlie LTP (Lyford et al., 1995, Guzowski et al., 2000).

Though attractive candidate genes the sustained nature of LTP suggests that a more prolonged response may be required to maintain potentiation than the relatively transient response directed by IEG. The mechanisms by which intracellular Ca^{2+} influxes, via the NMDA receptor, result in long lasting changes in synaptic efficiency are thought to involve second messenger pathways. One end-point of these signalling pathways is the phosphorylation of the cAMP response element binding protein (CREB) a constitutively active transcription factor (Rehfuss et al., 1991).

The finding that gene transcription and *de novo* protein synthesis are required for both l-LTP and LTF prompted Kandel to apply his findings, that LTF induction resulted in an increase in Cyclic-3',5'-Adenosine Monophosphate (cyclic AMP) levels (Cedar and Schwartz, 1972) to l-LTP. In neurons cAMP is generated from adenosine triphosphate (ATP) in direct response to increases in Ca^{2+} influxes from the NMDA receptor. The membrane bound enzymes adenylyl cyclase (AC) bind to calmodulin (Calm) only when Calm has bound Ca^{2+} (designated CaM). When CaM binds to AC, ATP is converted to cAMP (figure 1.10) (Xia and Storm, 1997). In the brain cAMP binds cAMP dependent protein kinase A (PKA). This binding results in dissociation of PKA into a regulatory dimer and two monomeric active sub-units (Taylor et al., 1990). The active catalytic subunit is capable of directly stimulating short term facilitation and is precisely regulated during LTF (Klein and Kandel, 1980). In attempting to understand the role of PKA in LTP transgenic mice expressing a mutated form of the regulatory region of PKA, that inhibits its activity, have been developed (Abel et al., 1997). The mutant mice display normal levels of neuronal activity and short-term LTP, however l-LTP is abolished. Similar results are obtained when PKA inhibitors are applied to hippocampal slices and if used in conjunction with a protein synthesis inhibitor. The coincident timing of the

effects of inhibition of protein synthesis and blocking of PKA led to the proposal that the main effect of PKA is to stimulate protein synthesis (Nguyen and Kandel, 1996). One of the downstream targets of activated PKA is CREB. Thus a possible link between Ca^{2+} signalling and gene transcription is via PKA signalling leading to CREB phosphorylation and activation. Research has identified two other signalling pathways by which Ca^{2+} increases can result in CREB activation.

Increased intracellular Ca^{2+} levels lead to the activation of the mitogen activated protein kinase or extracellular related kinase (MAPK or ERK) signalling cascade (figure 1.10), a core regulatory pathway of cell division and differentiation (Pearson et al., 2001). Inhibition of ERK suppresses LTP in the hippocampus and induction of LTP leads to rapid phosphorylation of ERK (English and Sweatt, 1997, Impey et al., 1998). The MAPK/ERK pathway starts with an intracellular increase of Ca^{2+}. Activation of the G-protein Ras, either by CaM, through CAMKII inhibition of SynGAP or tyrosine kinase activation, leads to the association of Ras with Raf and 14-3-3 proteins (Rosen et al., 1994). This association induces the phosphorylation of Raf which can also occur through cAMP signalling as opposed to CaM. Here, CaM induced generation of cAMP, via AC, activates Rap1 via cAMP GEF or interestingly through PKA (Kawasaki et al., 1998, Vossler et al., 1997). Thus, some of the effects of cAMP that have been attributed to PKA, may have been the result of ERK activation. Rap 1 then possibly activates Raf through phosphorylation. It is with Raf phosphorylation that the two pathways converge, leading to the sequential phosphorylation and activation of ERK via MEK (Impey et al., 1999). Activated ERK has a number of potential targets including cell adhesion molecules, cytoskeletal elements and ion channels. Thus activation of ERK could conceivably alter the postsynaptic membrane current or result in a conformational change within the dendrite (Sweatt, 2004). Importantly, ERK, when phosphorylated, translocates to the nucleus and brings about the phosphorylation and subsequent activation of CREB and a host of other transcription factors including c-Myc, c-fos, C/EBPβ and ATF-2 (Lynch, 2004). One of the long term effects therefore of ERK phosphorylation is a change in expression of target genes (Frodin and Gammeltoft, 1999, Thomson et al., 1999).

The CaM Kinase pathway (figure 1.10) consists of three related members; CaMKI and CamKIV and an upstream activator CaM Kinase Kinase (CaMKK). Each member of the cascade has a catalytic domain and a regulatory region containing an autoinhibitory and CaM binding domain. Binding of CaM alters the molecular conformation such that

substrate binding occurs activating the enzymes. CaMKI is further activated by phopshorylation (Frangakis et al., 1991,Cruzalegui et al., 1992) and while it is capable of activating CREB *in vitro*, if this occurs is uncertain, thus its key role is probably phosphorylation of CaMKK which in turn activates CaMKIV (Picciotto et al., 1995, Sun et al., 1996). CaMKIV shows strong nuclear localisation and evidence suggests it strongly activates CREB (Enslen et al., 1994, Matthews et al., 1994), and this is increased further in the presence of CaMKK. In addition CaMKIV may have a regulatory role in CRE- mediated transcription as it phosphorylates the CREB binding protein (CBP, see figure 1.11 , Chawla et al., 1998).

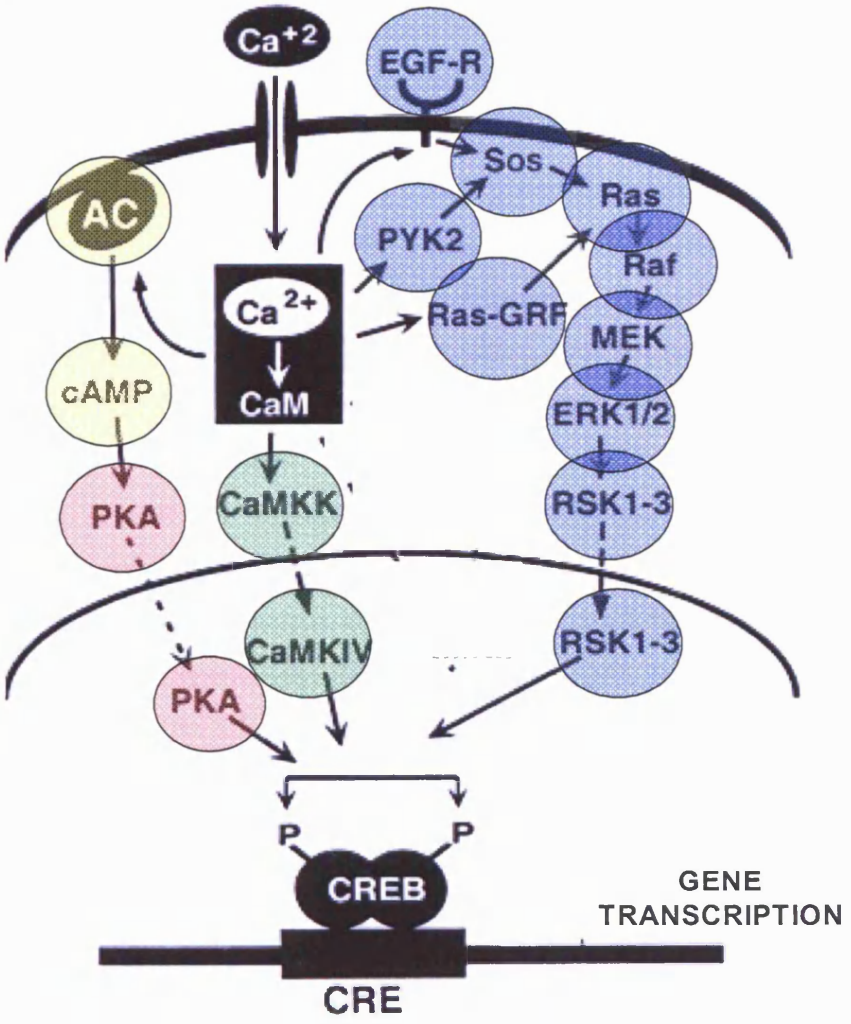


Figure 1.10 The signalling pathways that lead to CREB activation

CREB activation pathways are intitiated by the influx of Ca2+ into the postsynaptic neuron via NMDA channels. Ca2+ binds to Calmodulin (when bound is termed CaM). Either cAMP is generated by adenylyl cyclase activation by CaM, or CaM directly activates the CaM Kinase pathway (green circles).

cAMP activates both the PKA pathway (red circles) or the MAP Kinase pathway (blue circles). Adapted from West et al., 2001.

Thus a number of pathways could result in gene transcription through CREB phosphorylation following Ca2+ influx (figure 1.10). These pathways largely rely on sequential phosphorylation reactions and some convergence of the pathways has been proposed. *in vitro* studies suggest that CREB induced transcription of a reporter gene is synergistically enhanced by PKA and αCaMKII (Dash et al., 1991). Also, the translocation of MAPK to the nucleus, one requirement for CREB phosphorylation, requires PKA, suggesting a dual function of PKA via its own pathway or in concert with another (Impey et al., 1998). Further to this PKA is capable of phosphorylating two sites on CaMKK, inhibiting or enhancing CaMKK function, depending on the temporal sequence of stimulation ie if Ca2+ levels rise before cAMP is generated (Wayman et al., 1996, 1997). Cross-talk between CaMKIV and MAPK pathways can also occur through the CaMKIV phosphorylation of Rap and MAPK.

Finally the MAPK pathway produces slow CREB phosphorylation in contrast to the rapid action by CaMKIV (Finkbeiner et al., 1997) while CaMKIV inhibition disrupts slow and fast CREB phosphorylation suggesting a convergence of the two pathways (Wu et al., 2001). It has been found that CREB phosphorylation does not necessarily correlate with CREB-dependent gene expression, and correlation is dependent on the stimulus given to the cultured cells. In contrast to *Aplysia*, PKA inhibitors fail to block CREB phosphorylation in cultured hippocampal neurons subjected to synaptic stimulation whereas CaMKI inhibitors do (Deisseroth et al., 1996, Bito et al., 1996) In addition CaMKIV is observed in the nucleus at a similar time to CREB phosphorylation. This suggests that the signalling pathways dependent on the CaMK pathway may be crucial for the phosphorylation of CREB in cultured hippocampal neurons. In cultured hippocampal neurons CREB phosphorylated at Ser-133 can be detected using specific antibodies. Using this approach it was shown that the phosphorylation of CREB, induced by NMDA receptor synaptic stimulation, does not occur when an action potential is generated in the cell, suggesting that CREB phosphorylation is not a marker for activity but responds to specific Ca^{2+} signalling (Deisseroth et al., 1996).

1.3.5 CREB; TRANSCRIPTION AND NEURONAL PLASTICITY

The CREB family of transcription factors belong to the bZIP superfamily all of which contain a leucine zipper domain, which facilitates dimerisation, and a C-terminal basic domain that mediates DNA binding. The CREB family consists of CREB, cAMP

response element modulator (CREM) and activating transcription factors (ATFs). The bZIP domain shows high homology amongst all three members allowing the formation of hetero- and homodimers (Mayr and Montminy, 2001). CREB was first identified in experiments that searched for proteins that bind to the somatostatin promoter, specifically the palindromic sequence 5' TGACGTCA 3'. This sequence had been identified in the somatostatin promoter as necessary for its upregulation in response to elevated cAMP levels and PKA activation. Thus the sequence is commonly referred to as the CRE element (cAMP response element) (Montminy and Bilezikjian, 1987). CREB works in tandem with its co-activator CBP (CREB binding protein) to drive transcription (figure 1.11) (Chrivia et al.,1993).

It has been determined that CREB contains two glutamine rich domains, Q1 and Q2/CAD (constitutively active domain), separated by a kinase inducible domain (KID) (figure 1.11). In the KID domain at Ser-133 CREB is capable of being phosphorylated eg by kinases discussed more fully in section 1.3.4 (Kwok et al., 1994). The substitution of Ser-133 abolishes both CREB – CBP binding and phosphorylation dependent CREB mediated transcription (Gonzalez and Montminy, 1989), whereas mutations at other points in the domain do not affect CREB function. Thus, Ser-133 phosphorylation is considered necessary and, through loss-of-function mutant studies, sufficient for stimulus driven CREB-dependent gene expression. The constitutively active domain Q2/CAD, is necessary and sufficient for basal transcription as a low level of expression is driven even without stimulus (Quinn, 1993, Xing and Quinn, 1994). This is due to the fact that Q2/CAD associates with the basal transcriptional complex (Felinski and Quinn, 1999, Felinski et al., 2001). Thus Q2/CAD functions independently of KID though the reverse may not be true as Q2/CAD is necessary for the stimulus induced transcription that occurs at KID (Quinn 1993).

In an unstimulated state, therefore CREB uses its CAD domain to drive a basal level of transcription using the transcriptional machinery. On phosphorylation at Ser-133, CBP binds to CREB which serves two purposes. Firstly CBP is capable of binding the basal transcriptional machinery and therefore stabilises the whole complex at the CRE promoter. Secondly, CBP is known to have histone acetyltransferase activity which allows the dense chromatin to unravel and makes the complex more accessible for transcriptional molecules (Kwok et al., 1994, Nakajima et al., 1997, Bannister and Kouzarides, 1996).

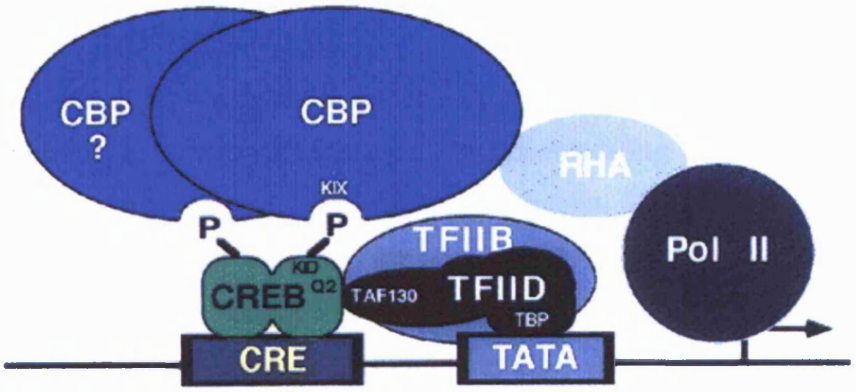


Figure 1.11 CREB and transcriptional interactions

Different domains of CREB bind distinct coactivators and basal transcription factors to activate transcription. When phosphorylated CREB dimerises and binds to the CRE element on the promoter of a target gene. TAF130 transcriptional element binds to the Q2 domain of CREB. A distinct domain of CREB, the KID, contributes to signal-induced transcriptional activation. When phosphorylated at Ser133, the KID of CREB can bind to the KIX domain of CBP (CREB Binding Protein). CBP with and stabilises the complex on the promoter site, whereas the Q2 domain of CREB interacts with other elements of the basal transcription machinery that are required for transcription, such as TFIID and TFIIB. Adapted from West et al., 2001.

The possibility that CREB activation is responsible for the gene expression that underlies neuronal plasticity has made it the target for a great deal of research in a number of plasticity paradigms The fact that CREB is required to invoke synaptic changes in most plasticity paradigms, highlights both the importance of gene transcription, and the key role CREB plays in regulating gene expression, in neuronal plasticity (Deisseroth et al., 1996 Mower et al., 2002).

In *Aplysia* the injection of oligonucleotides complementary to the CRE-promoter element, effectively blocks CREB function leading to a loss of LTF (Dash et al., 1990). The central role played by CREB in LTF, is further demonstrated by the fact that a cloned form of phosphorylated CREB, when injected into *Aplysia* induces LTF (Bartsch et al., 1995).

The development of transgenic technology has provided remarkable insights into the function of CREB in learning and memory. In 1994, Hummler et al., used gene targeting to disrupt exon 2 of the *Creb* gene in mice. The transgenic rodents failed to express the predominant CREB α / Δ isoforms (Hummler et al., 1994). In hippocampal slices $CREB^{\alpha}{}^{\Delta}$ mutants fail to display l-LTP though short term LTP was unaffected. The involvement

of CREB in experience dependent barrel cortex plasticity has been described (Glazewski et al., 1999). Following univibrissal deprivation the potentiation of neurons in surrounding barrels to the spared vibrissae, the marker of plasticity in the barrel cortex, is impaired in $CREB^{\alpha\Delta-}$ mutants

Barco et al., (2002) investigated the effects of a constitutively active form of CREB on LTP. Transgenic mice were developed by incorporating a viral protein under the inducible control of doxycycline into the mouse genome. Barco et al., used an expression vector encoding a CREB-VP16 fusion protein (CREB-VP16), which binds to cAMP response elements (CRE) through the basic-leucine zipper (B-ZIP) domain of CREB and activates transcription through the transcriptional activation domain of the herpes virus VP16 protein. The transgene was obtained by replacing the first transactivation domain of CREB with the acidic transactivation domain of the herpes virus VP16 protein, thus behaving like phosphorylated CREB. The findings were that a single tetanic train, usually only enough to generate short term LTP invoked l-LTP (Barco et al., 2002).

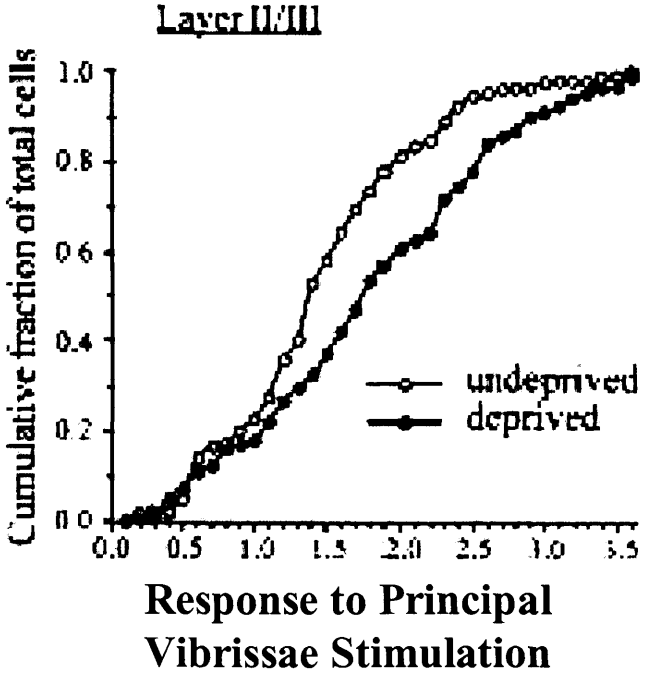


Figure 1.12 Potentiation of responses of cells in layer II/III following vibrissal deprivation

Experimental animals were deprived of all but one vibrissae for one day. Control animals were undeprived. Responses were recorded to stimulation of the principal vibrissae. The responses increase in deprived animals (closed circles) compared to controls (open circles) indicating potentiation of the cells has occurred following deprivation. Adapted from Barth et al., 2000

The well defined structure of the mouse barrel cortex has provided further evidence that CRE-mediated gene expression is associated with experience dependent plasticity. Using transgenic mice Barth et al., (2000) showed that significant CRE-mediated gene expression occurs in barrels that display plasticity following vibrissal deprivation. Specifically transgenic mice carrying a *lac Z reporter* gene downstream of six CRE elements were unilaterally subjected to one day deprivation of all but one vibrissae. Similar animals were subjected to either no deprivation, or all vibrissae were unilaterally deprived. When suitably processed the transgenic animals allow for *lac Z* expression to be visualised, thus the site of activated CREB can be inferred. While control and all deprived animals showed no significant *lac Z* expression in any place in the barrel field the univibrissal deprived animals showed significant upregulation in layer IV of the barrel representing the undeprived vibrissae (figure 1.13). Separately the effect of one days vibrissal deprivation on potentiation to the spared vibrissae was tested. Though approximately 14 days deprivation is required to induce complete plasticity (Glazewski and Fox, 1996) one days deprivation was sufficient to induce potentiation to the spared vibrissae in a subset of cells in layers II/III (figure 1.12).

This finding raised two important issues. Firstly it unambiguously showed that CRE-mediated gene expression is increased in the barrel cortex in response to patterns of vibrissal deprivation that induce neuronal plasticity and that the locus of plasticity is highly correlated with transcription. Secondly, as CRE-mediated gene expression occurs in layer IV but plasticity was observed only in layers II/III the findings suggested that CRE-mediated gene transcription plays a role in the pre-synaptic expression of plasticity (Barth et al., 2000).

1.3.6 HIGH THROUGHPUT TECHNIQUES FOR GENE EXPRESSION ANALYSIS

Research has provided important insights into the role various signalling pathways and CREB play in the gene transcription necessary for neuronal plasticity. Identifying those genes differentially expressed when plasticity is invoked will further understanding of the biological mechanisms that underlie plasticity in the barrel cortex. To date, identifying differentially expressed genes in the barrel cortex following vibrissal deprivation has relied on techniques that quantify expression of a single transcript for each biological sample. Here, *In Situ* hybridisations and RT-PCR have separately been used to detect differential expression of four immediate early genes and *Bdnf* in the barrel cortex following the induction of plasticity by vibrissal deprivation (Bisler et al., 2002, Nanda

and Mack, 2000). Morphological changes observed at the synapse following the induction of LTP include; increases in postsynaptic surface area, changes in the distribution and numbers of synaptic vesicles and synaptic curvature, increases in the numbers of dendritic spines, neuroskeletal rearrangements and changes to the extracellular matrix (Desmond and Levy, 1990, Applegate et al., 1987, Meshul and Hopkins, 1990, Nakic et al., 1998). It is likely that the structural changes that result in plasticity in the barrel cortex are dependent on similarly complex processes reliant on the co-ordinated actions of a large number of proteins, translated from a regulated network of gene transcripts. As roughly half of the estimated 22,000 genes in the mouse genome are expressed in the mammalian brain (www.ensembl.org/Mus_musculus/, Luo and Geschwind, 2001) identifying those associated with barrel cortex plasticity requires a high-throughput technique. High-throughput techniques provide a global view of gene expression and have the potential to unmask the functional interrelationships of plasticity related genes and hence the biological changes that result in plasticity.

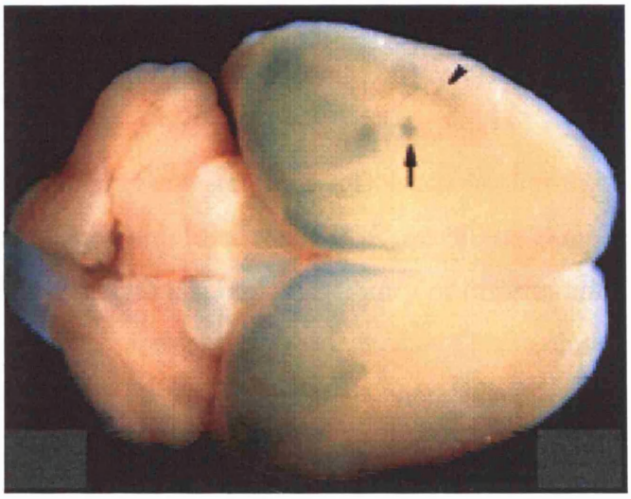


Figure 1.13 CRE-*Lac Z* expression in mouse brain following 16 hours vibrissal deprivation
Following 16 hours of unilateral single whisker deprivation whole brains were processed to visualise CRE-*lac Z* expression. Significant upregulation was observed in the precise location of the barrel corresponding to the deprived vibrissae (arrow). On the opposite hemisphere no significant *lac Z* expression is observed in the area of the barrel cortex. Adapted from Barth et al., 2000.

Differential display is a PCR based method that allows for the systematic comparison of transcripts expressed in various cells or tissues (Liang and Pardee, 1992) and has been

used to identify changes in gene expression in the rat hippocampus during LTP (Matsuo et al., 1998). The technique involves generating a cDNA representation of the transcripts in each sample. The cDNA is used as template in a PCR reaction, which uses an arbritary primer and poly A primer to direct amplification. The PCR products are size-separated on a standard sequencing gel and visualised by autoradiography. Differentially expressed genes can be identified on the gel image by bands that are visible in one sample and not in another. The respective band is then cut out and sequenced to reveal the identity of the regulated gene (Liang and Pardee, 1992). Differential display is an easy to use technique that can compare gene expression in as many samples as the gel allows from a small amount of starting RNA. It is, however, subject to a high false positive rate. Further potential problems include its bias to 3'- ends in the cDNA essentially meaning that definitive identification entails cDNA library screening. There are also sensitivity issues such as whether the technique can only detect genes that are absent from those that are present (Livesey and Hunt, 1996).

An alternative high throughput method of monitoring gene expression is the use of subtractive hybridisations (Diatchenko et al., 1996). Typically two cDNA sources are generated from RNA populations. Adaptors are ligated to a "tester" pool of cDNA usually the experimental sample. A second pool of cDNA "driver", usually the control sample, is added in excess and the two populations are co-hybridised. Due to the presence of the adaptors in the tester, those cDNA molecules that are common to both the tester and driver are not amplified in a PCR reaction. Amplified PCR products are cleaned and used to generate a subtractive library of differentially expressed genes. The library is then screened and clones sequenced. This technique requires only small amounts of RNA though its sensitivity limits it to detecting only highly differentially expressed genes. Feng et al., (2004) used subtractive hybridisation to identify genes differentially expressed in the visual cortex of rats during a period of plasticity and identified 30 genes significantly differentially expressed between the two samples.

Serial Analysis of Gene Expression (SAGE), in keeping with other methods generates cDNA from transcripts in RNA samples. Short tags are incorporated at the end of each cDNA molecule that allows them to be concatenated. The chain of cDNA molecules is then sequenced allowing the cataloguing of thousands of genes expressed in a sample, and importantly a quantitative measure of the number of copies of the same transcript Velculescu et al., 1995). Through specialised software multiple SAGE analysis comparisons can be run amongst various samples to identify differentially expressed

genes. Through database projects such as SAGEmap at the NCBI generated SAGE expression data can be compared with established libraries, including 20 human brain libraries, allowing global comparisons amongst, for example, various brain regions. The only disadvantage apparent with SAGE is the intensive sequencing facilities required to generate data (Colantuoni et al., 2000).

Cheng et al., (2003) have used SAGE to analyse the effects of monocular deprivation on gene expression in the extraocular muscle of the rat during dark rearing. Approximately 280 differentially expressed genes were identified of which 80 were characterised with a known function. Functional grouping suggested that the deprivation had resulted in a downregulation of genes involved in protein synthesis, lipid metabolism and muscle related processes.

Microarray technology has evolved from classic hybridisation techniques used to investigate specific transcript levels in RNA populations eg Southern and Northern blots (Southern, 1975). These techniques identified gene transcripts from within nucleic acid populations using a radioactively labelled nucleic acid probe complementary in sequence to the gene of interest. The nucleic acid population would be attached to a solid support and overlaid with the probe. Following hybridisation and washing, radioactivity within a sample was quantified providing a measure of the transcript level (Mataga et al., 2001). The principles of these techniques are reversed in microarray expression profiling. Nucleic acid molecules with sequences complementary to the transcripts being investigated are attached to a solid surface and all mRNA molecules within a population are labelled and overlaid onto the solid surface. This means that the number of specific gene transcripts being measured in a single RNA sample is limited only by the number of complementary sequences that can be attached to the solid matrix (Duggan et al., 1999). Following hybridisation the solid support is scanned and the amount of label at the probe is quantified, providing a relative measure of the amount of transcript in a sample. Nylon arrays, in which the RNA sample is usually radioactively labelled, have been largely superceded by glass arrays that use fluorescent technology (Schena et al., 1996, De Risi et al., 1997). Glass arrays have significant advantages over nylon arrays in that they allow for finer deposition of the target. This means that a greater number of genes can be assayed in a single experiment compared to nylon arrays.

Two forms of glass microarray are used in expression analyses. Oligonucleotide chips rely on in-situ synthesis of probe on the chip while cDNA chips are created by the robotically controlled deposition of DNA (Holloway et al., 2002). Due to the high costs

of manufacturing oligonucleotide arrays in house production facilities are rare. Companies such as Affymetrix (www.affymetrix.com, Santa Clara, CA.) or Agilent Technologies (www.agilent.com, Paolo Alto,CA) offer printed oligonucleotide chips containing up to 30,000 genes for a number of different organisms. Due to the high density afforded by short oligonucleotides each gene is represented many times on the chip by oligonucleotides complementary to different exons of the same gene thus giving the investigator greater confidence in signal detection above background noise. Each gene probe (perfect match, PM) is adjacent to a so-called mismatch (MM) probe which provides a control measure of non-specific background hybridisation. Quantifying the amount of hybridisation at the PM minus the value at the MM for each probe per gene allows the user to measure the amount of that gene in the RNA used (Holloway et al., 2002). By comparing the values obtained for the gene between two different RNA samples the user can tell if RNA 1 contains more of this gene than RNA 2. One drawback with oligonucleotide arrays is the lack of flexibility offered in the choice of genes spotted on each array. Oligonucleotide arrays tend to biased towards those organisms with well characterised genomes limiting their use for certain investigators. cDNA arrays by contrast have the capacity to print any source of cDNA onto the chip (Bowtell, 1999, Holloway et al., 2002). The cost of setting up a printing facility is also relatively cheap and due to the flexibility of the system a variety of projects involving many different organisms can take advantage of the equipment. The necessity to clone genes for spotting on cDNA arrays does add to the experimental input required and has the potential to introduce greater error into the system. Less steps are required to generate oligonucleotide arrays and as a consequence experimental error is reduced.

The flexibility afforded by cDNA array systems means that, in theory , any gene from any organism can be cloned and printed onto an array, the only limitation being the time and resources available to clone the required genes. This problem can be overcome by the availability of clone sets from various sources eg the IMAGE consortium (http://image.llnl.gov/) supply their clones free of charge to academic users, through licensed distributors (see http://www.hgmp.mrc.ac.uk/geneservice/index.shtml). Woo et al., (2004) directly compared the Affymetrix Gene Chip and cDNA microarray technologies with long oligonucleotide chips, which print 70mer sequences as opposed to manufacturing *in situ* (Hacia, 1999). The Affymetrix and long oligonucleotide chips show the best concordance, though the authors point out that no prior knowledge of the actual expression levels was available.

Both Affymetrix and cDNA technologies have been used to study gene expression in plasticity paradigms. Mayer et al., (2002) studied the effect of intense second messenger pathway activation on gene expression in cultured neurons (Mayer et al., 2002). By applying forskolin to the cells, cAMP production increases intracellularly mimicking the effect of receptor stimulation. The aim was to identify genes that may be differentially expressed under conditions that would trigger neuronal plasticity. The technology identified a number of significant genes that had been implicated in previous plasticity experiments in addition to a number of genes whose products are involved in signalling pathways themselves. A study that used cDNA technology following a subtractive hybridisation has identified genes involved in the spontaneous regeneration of the nervous system following injury. The study presented new hypotheses regarding the genes involved in the inhibition and promotion of spinal cord repair (Gris et al., 2003).

1.4 AIMS

A number of high throughput techniques for studying gene expression have been applied to various paradigms of plasticity though none to date have been applied to experience dependent plasticity in the barrel cortex. The aim of this thesis is to identify genes differentially expressed in the mouse barrel cortex under conditions that induce neuronal plasticity using microarray technology. In order to delineate changes in gene expression associated with plasticity from those associated with changes in activity levels as a result of vibrissal deprivation three forms of deprivation will be carried out. Whole barrel cortices will be extracted from animals whose vibrissae are either undeprived (control), all deprived (control for activity changes) or chessboard deprived (induces plasticity). RNA extracted from the barrel cortices will be used to generate fluorescently labelled probe and hybridised to two types of cDNA microarrays; a focused, custom microarray printed with transcripts of 370 candidate plasticity genes and a global microarray with 15,000 transcripts. Prior to conducting experiments the microarray process will be optimised. The technology will then be used to determine the relative amounts of transcripts represented on the microarrays in barrel cortices. Expression profiling will be conducted from early in the deprivation time scale (1 day) and at a number of time points until plasticity in the barrel cortex can be detected (8 days). At each time point the expression levels of transcripts in chessboard deprived, all deprived and undeprived vibrissae will be compared. As only chessboard deprivation induces plasticity statistical

techniques will be used to identify transcripts differentially expressed in chessboard compared to both deprived and undeprived barrel cortices.

CHAPTER 2

MATERIALS AND METHODS

2.1 ANIMAL HUSBANDRY

Animals were used in accordance with the guidelines set out in the Animals (Scientific Procedures) Act 1986. The programme of work was authorised in a Project Licence which stipulated the places of work and the Personal Licence holders qualified to carry it out.

2.2 GENETIC BACKGROUND OF EXPERIMENTAL ANIMALS

All mice used in this project were the male F1 offspring of crosses between C57/BL6 males (Harlan, Oxon, UK) and CBA/J females (Harlan, Oxon, UK), breeding was carried out in house. At the time of experimentation all animals were between 6-8 weeks in age.

2.3 VIBRISSAL DEPRIVATION

Vibrissal deprivation was carried out following the guidelines set out in Schedule 2a of the Animals (Scientific Procedures) Act 1986, (http://www.archive.official-documents.co.uk/document/hoc/321/321.htm). Animals were placed in a clear chamber and anaesthetised using a stream of isoflurane / O_2 (AstraZeneca, Cheshire, UK). Vibrissae were removed by applying slow tension to the base of the vibrissa using a forceps (Fox, 1992), a method which prevented damage to the follicle innervation (Li et al., 1995). Animals were allowed to recover before being returned to cages.

Two forms of deprivation were used. The first, referred to as DEP, involved removing all mystacial vibrissae in rows A, B, C ,D and E and the α, β, γ, and δ vibrissa (figure 2.1b). The second deprivation removed every other whisker, namely A2, A4, B1, B3, C2, C4, C6, D1, D3, D5, E2, E4, E6, in a chessboard (CB) pattern (figure 2.1c, Wallace and Fox 1992). Control, referred to as ALL, animals were subjected to brief anaesthesia and checked to ensure all vibrissae were present.

Under anaesthesia, animals were checked for vibrissal re-growth, every other day for the duration of the deprivation. Where necessary re-grown vibrissa were removed. Controls were also re-anaesthetised every other day for the period of deprivation. All deprivations

and checking for re-growth were carried out at the same time of day and the length of deprivation ranged from 0.4 - 8 days.

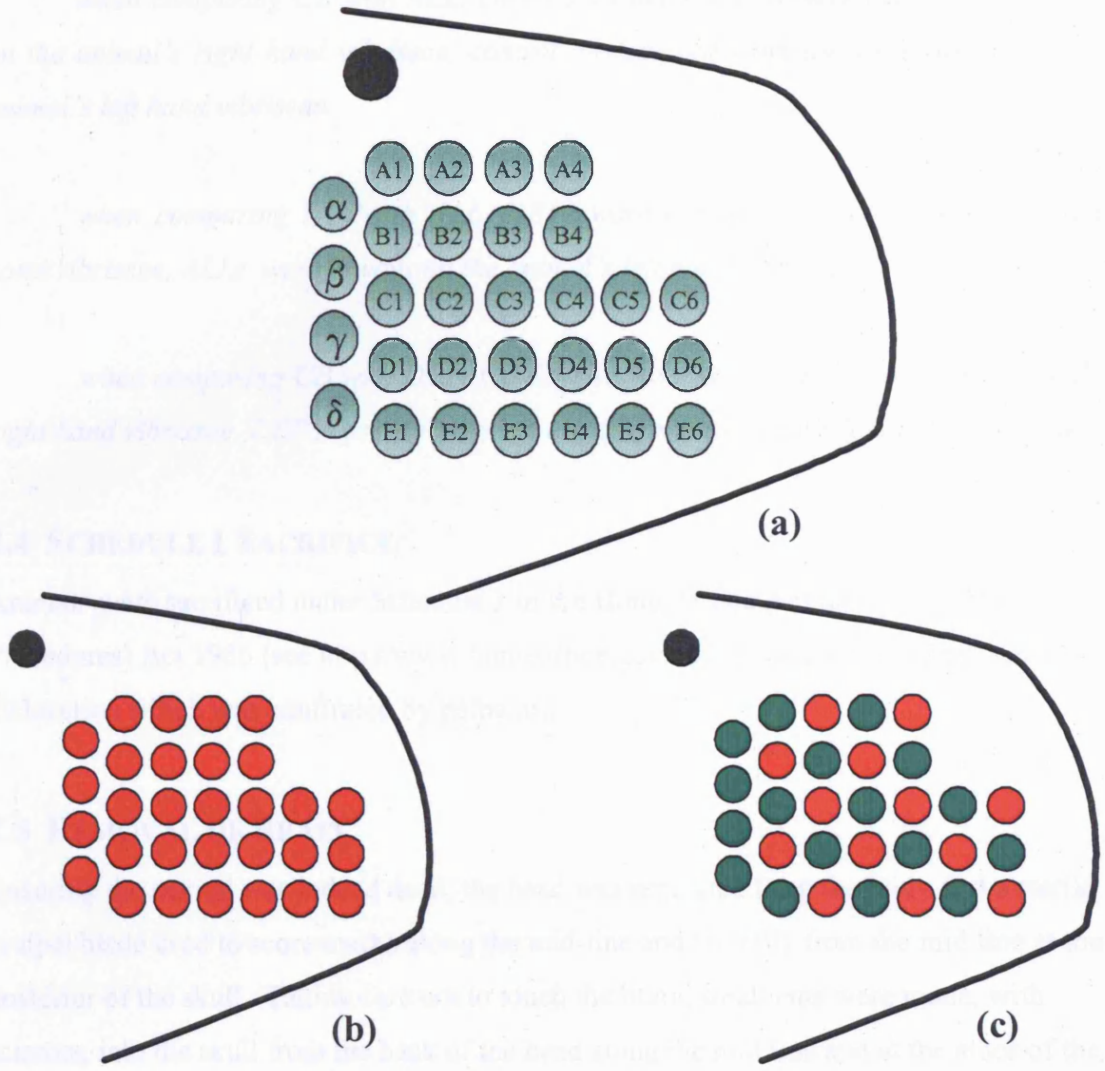


Figure 2.1 The patterns of vibrissal deprivations carried out on experimental animals.

(a) Undeprived vibrissae. Each circle represents a single vibrissa. The vibrissae are arranged in rows referred to as A-E. The numbers reflect the position of a single vibrissa within each row, vibrissa 1 being the most posterior. The vibrissae labelled α, β, γ and δ are considered as single entities and are not affiliated to any row. The green colour reflects the fact that vibrissae are spared (undeprived). In this example all vibrissae are spared providing a control barrel field.

(b) All deprived vibrissae. Red circles show position of deprived vibrissae. The identites of each individual vibrissa use the convention described in figure 2.a.

(c) The chessboard deprivation. Deprived vibrissae are coloured red, green circles indicate spared vibrissae. The identites of each individual vibrissa are given in figure 2.a.

In Section 4, two forms of deprivation are applied to single animals. To avoid confusion when processing samples, each animal was deprived using the following left / right hand side conventions:

when comparing CB with ALL, chessboard deprivations were always carried out on the animal's right hand vibrissae, control, undeprived vibrissae were always on the animal's left hand vibrissae.

when comparing DEPwith ALL, DEPs were carried out on the animal's right hand vibrissae, ALLs were always on the animal's left hand vibrissae.

when comparing CB with DEP, CB deprivations were carried out on the animal's right hand vibrissae, DEP deprivations were always on the animal's left hand vibrissae.

2.4 SCHEDULE 1 SACRIFICE

Animals were sacrificed under Schedule 1 of the Home Office Animals (Scientific Procedures) Act 1986 (see http://www.homeoffice.gov.uk/comrace/animals) by cervical dislocation which was confirmed by palpation.

2.5 REMOVAL OF BRAIN

Ensuring the animal was indeed dead, the head was separated from the body and a sterile scalpel blade used to score marks along the mid-line and laterally from the mid line at the posterior of the skull. Taking care not to touch the brain, small cuts were made, with scissors, into the skull from the back of the head along the mid line and at the sides of the skull. Small forceps were used to prise apart and remove each half of the skull exposing the brain. A small spatula was inserted under the remaining skull at the anterior of the brain and the brain removed.

2.6 BARREL CORTEX EXTRACTION

2.6.1 RIG DESIGN

Figure 2.2 is a photograph of the rig designed to extract the barrel field. In the solid base is a plaster mould (M) of the underside of a mouse brain. This held the brain in place while manipulations were carried out. The supporting pin (P) holds the adjustable X-Bar

(X) in place. Attached to the X-Bar is an adjustable Y-Bar (Y) which was connected to the borer guide (G). A 4 mm hole was drilled in the borer guide which allowed a stainless steel borer (B) to be held in place. The borer had a 2mm diameter hole with a sharpened cutting edge.

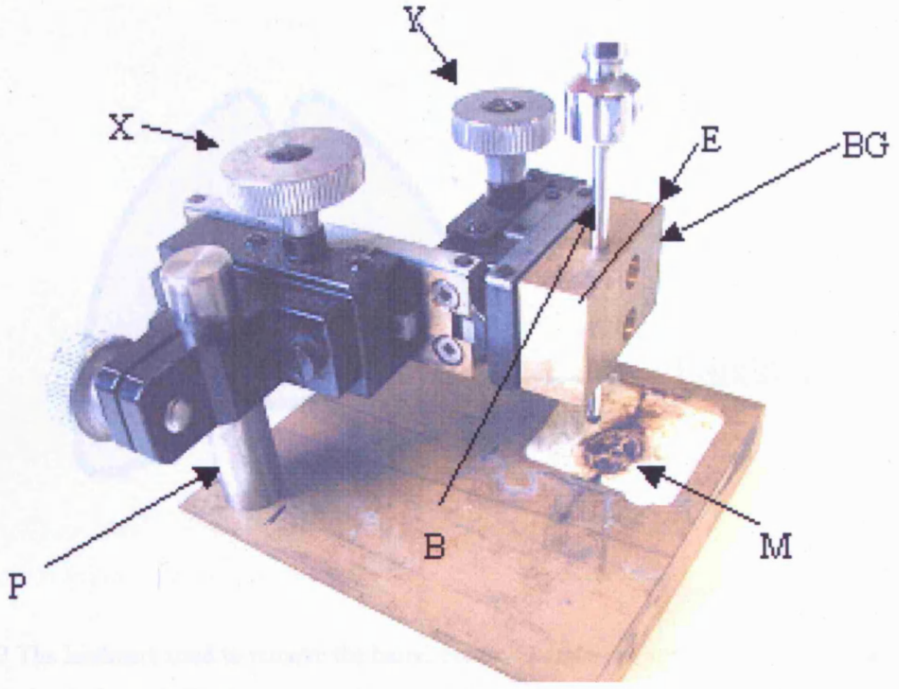


Figure 2.2 Rig for extraction of the mouse barrel cortex
A mouse brain sits in the mould, M. The supporting pin (P) allows adjustment of the edge of (E) of the borer guide (BG), such that it lies parallel to the mid-line of the brain. Adjusters, X and Y, allow the required lateral and anterior / posterior movement of BG (correct to 0.1mm). A borer is then used to sample the relevant area of cortical tissue.

2.6.2 USING THE RIG TO REMOVE THE BARREL CORTEX

The primary use of the rig was in the removal of the mouse barrel cortex. A suitably intact brain (see section 2.5) was first placed in the mould (figure 2.2). By loosening the attachment to the supporting pin, the borer guide could be adjusted so that it's edge (E, figure 2.2) ran parallel to the mid line of the brain. The borer was then placed through the guide and the X- and Y- bars adjusted such that the cutting hole of the borer was centered above a ***landmark*** present on every mouse brain (see figure 2.3). The borer was gently pressed against cortex to ensure accurate placement of borer with respect to ***landmark*** (a dot was visible in the centre of the circle made by borer (figure 2.3).

The barrel cortex was removed by moving the borer guide 2.1mm anterior to the landmark and 2.8mm laterally (either left or right). This placed the borer in the precise location of the barrel cortex. With gentle pressure the borer was used to cut through the cortex allowing the tissue to be extracted with a forceps.

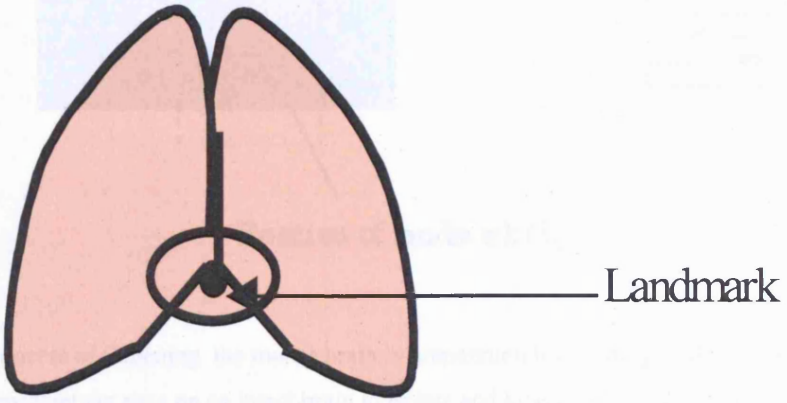


Figure 2.3 The landmark used to remove the barrel cortex. ***Landmark*** appears as a dot in the centre of the circle made by the borer indicating correct alignment of the brain with respect to the borer.

2.7 CYTOCHROME OXIDASE STAINING OF SECTIONED CORTICAL TISSUE

2.7.1 PREPARATION OF CORTICAL TISSUE FOR MICROTOME SECTIONING

Brains were removed (section 2.5) and placed in fixative (see table 2.1) overnight in a petri dish. The following day the brain was removed and a single edge razor blade was used to remove the cerebellum and approximately 3mm of the anterior of the cortex. The brain was divided along the mid line (see figure 2.4 a). The bottom of each hemisphere was cut at an angle of approximately 70° (see figure 2.4 b). The respective hemispheres were then flattened between 2 glass microscope slides (Fisher Scientific, Loughborough, UK), held together with plasticine (Sigma, UK) and returned to fixative (figure 2.4 c). This served to place the barrel field in the horizontal plane so that it could be viewed as a complete entity when sectioned and stained.

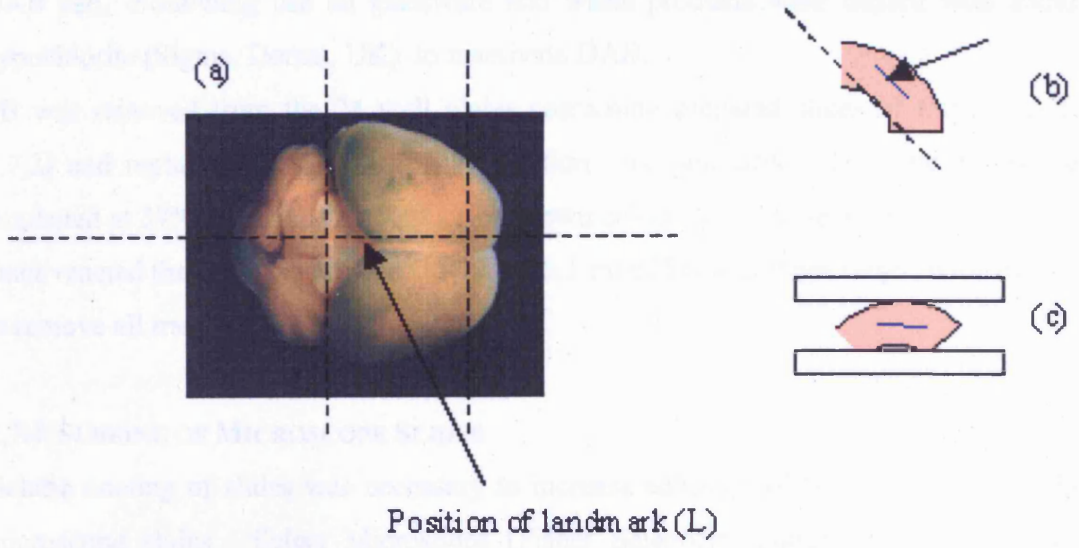


Figure 2.4 The process of flattening the mouse brain in preparation for staining of the mouse barrel cortex.
a) dotted lines represent cut sites on an intact brain to isolate and bisect the cortex.
b) a single hemisphere is cut at approx. 70° (frontal view of cortex shown) such that flattening the cortex (orange) between two slides, (c), results in the barrel field (blue line) being in the horizontal plane.

2.7.2 SECTIONING OF CORTICAL TISSUE

Following overnight fixation the flattened cortex was sectioned in 35µm layers using a Leica SM2000R microtome (Leica Microsystems, Bucks., UK) with an attached freezing stage. A flat cutting stage was formed, by adding 20%w/v sucrose in phosphate buffer (PB, table 2.1), to the stage at approx –20°C. The frozen solution was shaved down until a flat surface, large enough to accomodate the tissue was available. The flattened cortex was placed on the surface and while being frozen was flattened with a slide. Once frozen the tissue was covered with 20%w/v sucrose in PB which froze covering the tissue. 35µm sections were cut and transferred to individual wells of a 24 well culture plate (Corning, NY, USA), containing pre-heated, at 37°C, PB (no sucrose). In order to remove all traces of fixative the slices were rinsed three times in fresh PB prior to further work.

2.7.3 CYTOCHROME OXIDASE STAINING OF CORTICAL TISSUE

The staining protocol made use of diaminobenzidine (DAB, Sigma, Dorset, UK), a potent carcinogen. Extreme care was taken when using DAB and all work was carried out in designated areas to minimise hazards, including fume hoods and ovens specifically for

DAB use. Following use all glassware and waste products were treated with sodium hypochlorite (Sigma, Dorset, UK) to inactivate DAB.

PB was removed from the 24 well plates containing prepared slices of tissue (section 2.7.2) and replaced with 1ml of DAB reaction mix (see table 2.1). The plates were incubated at 37°C until slices turned a rich brown colour (usually between 3 and 5 hours). Once reacted the DAB was replaced with with 1 ml of PB and slices rinsed with fresh PB to remove all traces of DAB.

2.7.4 SUBBING OF MICROSCOPE SLIDES

Gelatin coating of slides was necessary to increase adhesion of tissue sections to glass microscope slides. Select Microslides (Fisher Scientific, Loughborough, UK) were placed into a slide rack and transferred to a glass bath in which they were rinsed with deionised H_2O. The slides were immersed in subbing solution (see table 2.1) for 1 minute, removed from the bath and dried overnight at 37°C.

2.7.5 MOUNTING OF TISSUE SECTIONS TO SLIDES

Required tissue sections were carefully transferred to subbed slides (section 2.7.4) using a fine paint brush. The slides were dried out overnight at room teperature, then dipped in xylene (VWR™ International, Dorset, UK) for 1 minute. Excess xylene was removed by tissue blotting and the sections were covered with DPX mounting medium (VWR™ International, Dorset, UK). Carefully, to prevent air bubbles forming, the slides were overlaid with a cover slip (VWR™ International, Dorset, UK). On drying the slides were examined at the relevant magnification using a Leica DM2500 microscope (Leica Microsystems, Glattbrugg, Switzerland).

2.8 RNA PROTOCOLS

2.8.1 GENERAL CONSIDERATIONS FOR WORKING WITH RNA

RNA is extremely sensitive to degradation by RNAse nucleases which are endogenous to all cell types. As a result, all possible steps were taken to maintain an RNAse free working enviroment. Work surfaces were routinely wiped down with a 0.05M NaOH solution. Pipettes, forceps, scalpels and all other equipment that came into contact with RNA were treated similarly. If NaOH treatment was not possible eg Pipette tips,

eppendorf tubes, then autoclaving was carried out using a Prestige Medical 2100 benchtop autoclave (Prestige Medical, Lancs., UK) followed by drying in a dedicated oven. Autoclaving was regularly carried out, even on equipment claimed RNAse free by the manufacturer. These procedures were followed in addition to standard lab practices such as wearing gloves at all times and using fresh consumables such as pipette tips to avoid contamination. RNA was stored in RNAse free water (Ambion, Cambs., UK) at –80°C when storage was for a couple of weeks. For longer term storage 1/10 volume of sodium acetate and 2.5 volumes of ethanol were added to the solution prior to storage at –80°C.

2.8.2 TOTAL RNA EXTRACTION

Extraction of RNA was performed using TRI®Reagent (Sigma, Dorset, UK). TRI®Reagent isolation of RNA is a single step method developed by Chomcsynski (Chomcsynski and Sacchi 1987) which allows the rapid and cost-efficient purification of RNA from a number of sources.

Up to 100mg of tissue was placed in a sterile Bijou bottle (Sterilin, Staffs., UK) containing 1 ml of TRI®Reagent. The tissue was homogenized to disrupt cells using a Ultraturrax T25 homogenizor (IKA Labortechnik, Staufen, Germany). Prior to, after and between each homogenization the probe was sequentially washed with 4M NaOH, autoclaved Milli-Q H_2O and 95-100% ethanol. The homogenate was transferred to a 1.5ml eppendorf tube (Greiner Bio-One, Frickenhausen, Germany) and centrifuged for 10 minutes to pellet insoluble products such as fats and other cell debris. All centrifugation steps were carried out at 4°C using a Beckman Coulter Allegra®21R Centrifuge (Beckman Coulter®, Bucks, UK). The supernatant was transferred to a fresh ependorff and 200ul of chloroform (Fisher Scientific, Loughborough, UK) added, followed by vigourous shaking of the tube for 15 secs. The solution was allowed to stand for 15 minutes at room temperature then centrifuged for 15 minutes which seperated the mixture into three phases.

The upper aqueous phase contained RNA and this was carefully pipetted off and transferred to a fresh eppendorf (the interphase contains DNA and the bottom organic phase contained protein). Propan-2-ol (Fisher Scientific, Loughborough, UK) was added (500μl) to the aqueous phase, the mixture vortexed and left to stand at room temperature for 10 minutes. The precipitated RNA was pelleted by centrifugation for 10 minutes. The liquid was removed by pipetting, taking care not to disturb the pellet, which was then

washed with a 75% solution of ethanol (Fisher Scientific, Loughborough, UK). The washed RNA was then centrifuged, for 7 minutes, to pellet and the liquid removed by pipetting. After removing the majority of the liquid, tubes were briefly centrifuged and a pipette used to remove any last traces of ethanol. The pellets were air dried for approximately 5 minutes (taking care not to overdry) and resuspended in RNAse free water (Ambion, Cambs., UK).

2.8.3 PRECIPITATION / CONCENTRATION OF RNA

Sodium acetate, 1/10 volume of the RNA sample, and 2.5 volumes of ethanol were added to the RNA sample. The solution was mixed and stored overnight at -20°C. The following day samples were centrifuged for 15 minutes and the supernatant removed (as section 2.8.2). Pellets were washed with 70% ethanol, centrifuged for 7 minutes and the supernatant removed. The pellets were air dried and redissolved in the required volume of RNAse free H_2O (Ambion, Cambs., UK).

2.8.4 DNASE TREATMENT OF RNA

To remove contaminating DNA from RNA samples, 1 unit of RQ1 RNAse free DNAse (Promega, Southampton, UK) per ug of RNA, was added to the RNA solution and buffered using RQ1 10X buffer (Promega, Southampton, UK). The reaction was incubated at 37°C for 30 minutes and terminated by the addition of RQ1 stop solution (Promega, Southampton, UK). DNAse was inactivated by 15 minute incubation at 65°C.

2.8.5 RNA CLEAN-UP

Qiagen RNeasy columns were utilised to remove contaminants in RNA, carried over from TRI®Reagent extraction and DNAse treatment, using the manufacturer's protocol (Qiagen, West Sussex, UK).

2.8.6 GENERATION OF ANTI-SENSE RNA USING AN IN-VITRO TRANSCRIPTION REACTION

MessageAmp™ aRNA kit (Ambion, Cambs, UK) was used to amplify large amounts of anti-sense RNA (aRNA) from limited amounts of mouse barrel cortex total RNA. Prior to the amplification reaction , all total RNA samples were subjected to strict QC measures (section 3.3.2). The protocol is based on the procedure described by Eberwine and can be broken down into 5 steps (Van Gelder et al., 1990). Steps 1, 2 and 3 were carried out in

one day, steps 4 and 5 on a separate day. All procedures were carried out under strict RNAse free conditions (section 2.8.1).

Step 1 – First Strand cDNA Synthesis.

Aliquots (typically 1µg) of total RNA in 1.5ml Eppendorf tubes, were removed from -80°C storage and the volume adjusted to 11µl with RNAse free water (supplied in kit). To each sample 1µl of T7 oligo (dT) primer was added (supplied in kit), the tubes were spun briefly (5 seconds) and heated in a water bath at 70°C for 10 minutes. After snap cooling on ice, 8µl of first strand reaction mix was added to each sample. The reaction mix consisted of 2µl 10X First Strand Buffer, 1µl Ribonuclease Inhibitor, 4µl dNTP mix and 1µl Reverse Transcriptase (all included in kit). The reactions were mixed well by pipetting and spun briefly prior to incubation at 42°C for 2 hours in an oven.

Step 2 – Second Strand cDNA Synthesis.

Immediately following the first strand synthesis each sample was transferred to a 0.2ml PCR tube (ABgene®, UK) and placed on ice. Second strand cDNA reaction mix (80µl), consisting of RNAse free water (63µl), 10X Second Strand Buffer (10µl), dNTP mix (4µl), DNA polymerase (2µl) and 1µl RNase H was added to each sample. All reagents were supplied in the kit. Following mixing and brief centrifugation the samples were incubated at 16°C for 2 hours in a Flexigene thermocycler.

Step 3 – cDNA Purification.

cDNA generated during steps 1 & 2 were purified using microcentrifuge spin columns included in the kit. Manufacturer's protocols were followed and the purified cDNA samples stored at -20°C.

Step 4 – aRNA Synthesis.

To 16µl of the cDNA samples was added 4µl of each of the following: T7 ATP, T7 CTP, T7 GTP, T7 UTP, T7 10X Reaction Buffer and T7 Enzyme Mix (supplied). The samples were mixed by pipetting and briefly centrifuged, prior to a 12 hour incubation at 37°C. The incubations were carried out in an oven to reduce condensation preventing changes in concentration that would affect yield. Following incubation 2µl DNAse 1 (supplied) was added to each tube and incubation resumed for 30 minutes. This DNAse treatment removed all cDNA template.

Step 5 – aRNA Purification.

aRNA samples generated during step 4 were purified using microcentrifuge spin columns included in the kit. Manufacturer's protocols were followed and the purified aRNA samples stored at -80°C.

Agarose gel electrophoresis and UV spectrophotometry were used to assess aRNA integrity and purity (section 3.3.2).

2.8.7 REVERSE TRANSCRIPTION TO GENERATE cDNA

Total RNA (1ug), oligo dT primer (125ng) and random hexamers (250ng, both MWG, Germany), in 12ul RNAse free H_2O (Ambion, Cambs., UK), were placed in an autoclaved 1.5ml Eppendorf tube and heated at 70°C for 10 minutes then placed on ice. A first strand synthesis reaction mix, consisting of 1x MMLV buffer (see table 2.1), 0.5mM dNTP's, 10mM DTT (all final concentrations), 20U of RNAsin® and 400U of MMLV reverse transcriptase (all Promega, Southampton, UK) were added. The solution was mixed gently by pipetting, centrifuged briefly and incubated at 37°C for 50 minutes. The reaction was terminated by heating at 80°C for 10 minutes. Generated cDNA was stored at -20°C.

2.8.8 CREATION OF REFERENCE RNA FOR MICROARRAY HYBRIDISATIONS

The microarray hybridisations carried out in Chapter 5 required the use of a reference RNA sample which was generated by pooling approximately 0.3μg of total RNA from each experimental sample (section 5.5.5). Pooled total RNA was used to generate aRNA. Products of individual amplification reactions were then pooled to form an aRNA reference pool (section 2.8.6)

2.9 DNA PROTOCOLS

2.9.1 GENERAL CONSIDERATIONS

DNA is less susceptible to degradation, by nucleases, than RNA, nonetheless care was taken to prevent it, and contamination by using dedicated equipment and work spaces. In addition reagents were kept only for a minimum period of time, pipette tips were certified DNAse free and DNA was stored at -20°C.

2.9.2 POLYMERASE CHAIN REACTION

The polymerase chain reaction was used to amplify DNA specific sequences, directed by primers, using the enzyme Taq Polymerase (Mullis and Faloona, 1987).

2.9.2.1 Primer Design

Primer design was carried out using the Primer 3 on-line design software, developed by the Whitehead Institute and the Howard Hughes Medical Institute and located at http://www-genome.wi.mit.edu/genome_software/other/primer3.html . All primers were designed to incorporate a GC clamp at the 5' end, show no internal pairing, minimise dimer fromation and have an annealing temperature of approximately 58°C. All primers were purchased from MWG (MWG Biotech AG, Milton Keynes, UK) and were HPLC purifed. A complete list of primer sequences used is presented in Table 2.2.

2.9.2.2 PCR Reaction Mix

All PCR reactions were carried out in either 200 μl thin-walled PCR tubes (ABgene, Epsom, UK) or 96 well PCR reaction skirted plates (ABgene, Epsom, UK). The PCR reaction mixture contained the target DNA template (1-2 μl), 3 units Taq polymerase (Promega, Southampton, UK), 1 μM of forward and reverse oligonucleotide primers and 0.2 mM dNTPs from 10mM stocks of dGTP, dCTP, dTTP and dATP (Promega, Southampton, UK). The reaction was bffered using a 1 x Taq polymerase reaction buffer (Promega, Southampton, UK, see table 2.1).

2.9.2.3 PCR Cycling Parameters

PCR reactions were performed in a Techne Flexigene thermocycler (Techne, Cambridge, UK) using the following cycling parameters:

2.9.3 SINGLE PCR REACTION CLEAN-UP

Where necessary PCR products where purified by QIAquick, PCR (Qiagen, West Sussex, UK) clean up columns following the manufacturer's protocol. The columns contained a silica-gel membrane that allowed PCR generated fragments to bind while washing off the remaining components of the reaction eg primers, dNTP's. The purified PCR products were eluted in the required volume of H_2O and stored at -20°C.

95°C	for	5 minutes (initial denaturation)	
95°C	for	30 seconds (chain melting)	x 30
58°C*	for	30 seconds (primer annealing)	
72°C	for	45 seconds (chain extension)	
72°C	for	5 minutes (final extension)	

* For some sets of primer pairs primer annealing temperature (Tm) and MgCl2 concentration needed optimisation to remove multiple bands / promote product formation. Optimal conditions for each primer pair are shown in table 2.2.

2.9.4 MULTIPLE PCR REACTION CLEAN-UP

A system was developed that allowed the purification of up to 96 PCR products in a single run. An entire bottle of Sephacryl S-300 (Sigma, Dorset, UK) was equilibrated by removing the ethanol and resuspending in 10mM Sodium Phosphate. The Sephacryl S-300 was then aliquoted into each well, 350µl/well, of a 96-well Silent Screen filter plate (Fisher Scientific, Loughborough, UK).

The Silent Screen plate was placed on top of a 96 well plate (ABgene, Epsom, UK) and centrifuged at 2000rpm for 2 minutes to remove all residual liquid. The Silent Screen plate was then placed on top of a fresh 96 well plate, and the relevant PCR product transferred to the wells, containing the Sephacryl S-300, taking care to pipette onto the centre of the Sephacryl. The plates were centrifuged at 2000rpm for 3 minutes, and the eluent containing the purified PCR product was retained. Purified cDNA was stored at -20°C.

2.9.5 DNA SEQUENCING

In certain instances, for example following cloning of DNA fragments (section 2.14), sequencing of the desired fragment was carried out to ensure the correct sequence had been inserted.

A reaction mix was prepared by combining 4µl of PCR generated sequence from the relevant clone, 1.6 pmol of primer (either forward or reverse, specific to the sequence of interest) and 4µl of Big Dye® V2 Mix (Applied Biosystems, CA, USA) with H_2O to a final volume of 10µl in a 0.2ml PCR tube (Applied Biosystems, CA, USA). The reaction

mix was placed in a Techne Flexigene thermocycler and subjected to the following thermal cycling conditions:

95°C	for	10sec	(chain melting)	
50°C	for	5 seconds	(primer annealing)	x 25
60°C	for	4 minutes	(extension)	
4°C			indefinitely (soak)	

Excess Big Dye V2 was removed from the reaction by a precipitation step prior to sequence analysis. Isopropanol (90µl) 70% (v/v Fisher Scientific, Loughborough, UK) was added, the samples were vortexed and the DNA was allowed to precipitate in the dark (20 minutes at room temperature). Samples were centrifuged at room temperature for 30minutes at 13,000 rpm. The isopropanol was removed, and the pellet rinsed with a further 150µl of 70% (v/v) isopropanol (Fisher Scientific, Loughborough, UK). Following vortexing, samples were again spun at RT for 5 minutes at 13,000rpm and the supernatant removed. Pellets were allowed to air-dry in the dark and the tubes wrapped in foil before storage at -20°C, prior to sequencing.

The resulting amplified DNA fragments, now fluorescently labelled, were dissolved in 15µl sequencing grade formamide (Fisher Scientific, Loughborough, UK) and denatured at 95°C for 2 minutes before placing on ice for 1 minute. Sequence analysis was performed by capillary electrophoresis using an ABI Prism® 3100 Genetic Analyzer (Cheshire, UK).

Sequences were analysed using BioEdit Sequence Alignment Program (available for download at http://www.mbio.ncsu.edu/BioEdit/bioedit.html) . Significant homology with the desired sequence was determined using the Blast 2 Sequences program at the NCBI (http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html).

2.10 AGAROSE GEL ELECTROPHORESIS

2.10.1 GEL ELECTROPHORESIS OF RNA / DNA SAMPLES

DNA/RNA samples were size fractionated by gel electrophoresis. Gels were prepared by melting agarose (Invitrogen™, Paisley, UK) in 1X TAE buffer (Invitrogen™, Paisley, UK), typically 1-2% w/v. On cooling (below 50 °C) ethidium bromide (Sigma, Dorset, UK) was added to a final concentration of 4ug/l and the gel allowed to set in an appropriate mould. The required amount of DNA/RNA was added to one fifth volume of loading dye (50% (v/v) glycerol and1% (w/v) bromophenol blue) and loaded into the set gel. 3µl of ϕ size ladder (Promega, Southampton, UK) were routinely added to all gels to gauge product size. Electrophoresis was performed at 110V for the required amount of time in a Pharmacia GNA-100 tank (Amersham Biosciences, Bucks., UK) in 1X TAE buffer. Nucleic acid bands were visualised under UV light using a Syngene Gene Genius Bioimaging System (Syngene, Cambridge, UK).

For RNA samples buffers and glassware were autoclaved and the gel tank, mould and combs were soaked in 0.1% solution SDS (Sigma, Dorset, UK) overnight prior to electrophoresis to prevent RNAse degradation.

2.10.2 HIGH THROUGHPUT GEL ELECTROPHORESIS

In situations where large numbers of samples needed to be visualised specialised gel moulds were used that allowed sample loading using a multi-channel pipette. A Flowgen FlowMadge® diagonal electrophoresis (Flowgen Bioscience Ltd., Notts., UK) or an ABgene® Electro-Fast® Stretch Gel System (ABgene, Epsom, UK) were used enabling up to 96 samples to be run simultaneously.

2.10.3 GEL ELECTROPHORESIS OF FLUORESCENTLY LABELLED CDNA

Gel electrophoresis was used to qualitatively assess the amount of fluorophore incorporated into cDNA generated following RNA labelling reactions (section 2.12). A reduced volume gel mould the size of a glass microscope slide was was filled with 1ml agarose (prepared as in section 2.6.1without ethidium bromide) and overlaid with a microscope slide. Once set, the gel was prised out and 1μl of sample (combined with 1μl glycerol, Sigma, Dorset, UK) was loaded into 1mm^2 wells and electrophoresed at 100V for approximately 20 minutes. Gels were then scanned, using a GeneTAC™LS-IV

scanner (Genomic Solutions, Cambs, UK), at 550nm and 650nm wavelength for Cy3 and Cy5 fluorophores respectively (section 2.17).

2.11 QUANTIFICATION OF NUCLEIC ACIDS

The absorbance of UV light, of 260nm and 280nm wavelength, by the nucleic acid sample was measured using a GeneQuant spectrophotometer (Amersham Biosciences, Bucks., UK). The spectrophotometer converted this measurement for RNA, or DNA into a concentration and calculated the OD260/280 ratio.

2.12 PRODUCTION OF FLUORESCENTLY LABELLED PROBE

Cyanine3 (Cy3) and Cyanine5 (Cy5) labelled dUTP (Amersham Biosciences, Bucks., UK) were used in all labelling reactions to produce fluorescently labelled probe prior to microarray hybridisations. Both fluorophores are extremely sensitive to light and care was taken to minimise their exposure by performing all reactions in sterile amber Eppendorf tubes (Fisher Scientific, Loughborough, UK). Multiple aliquots were taken from the stock solution to prevent freeze thaw effects.

2.12.1.1 Evaluation of Kits to Produce Fluorescently Labelled Probe

Aliqouts (10μg) from a total RNA pool (section 3.3.3) were used to compare the efficiency of 4 different methods to form fluorescently labelled probe (section 3.3.5.3). Commercially available kits were compared to a modified reverse transcription reaction.

2.12.1.2 Modified Reverse Transcription

The reaction was carried out as in section 2.8.7 except that the dNTP composition was adjusted to account for Cy labelled dUTP. dNTP mix had a final concentration of 0.5mM dATP, dGTP, dCTP and 0.1mM dTTP. Cy3/Cy5 labelled dUTPs (1mM stock, Amersham Biosciences, Bucks., UK) were added individually to each reaction to a final concentration of 0.02mM. Following each labelling reaction, removal of short degraded RNA oligomers and unincorporated nucleotides, was achieved using AutoSeq™ G-50 columns (Amersham Biosciences, Bucks., UK). Briefly, the column caps were unscrewed a quarter of a turn, the bottom snapped off and the column placed into a 1.5ml Eppendorf tube. The column was spun for 1minute at 2000 x g in a Beckman Coulter Allegra®21R Centrifuge (Bucks, UK). Excess fluid was blotted from the column tip and

the column transferred to a fresh amber Eppendorf tube. The reaction mix was pipetted onto the centre of the resin and the column respun. Purified cDNA was stored at –20°C.

2.12.1.3 CyScribe™ First-Strand cDNA Labelling Kit

CyScribe™ First-Strand cDNA Labelling Kit was purchased from Amersham Biosciences, Bucks., UK. The kit was supplemented with Cy3/Cy5 labelled dUTPs (Amersham Biosciences, Bucks., UK). Random hexamer and oligo dT primers (supplied) were used to prime total RNA template. Where aRNA (section 3.5.5.4) was used as template, only random hexamers were used in cDNA synthesis. Manufacturer's instructions were followed throughout the labelling reaction.

On completion of the labelling reaction any residual RNA was removed by adding NaOH to a final concentration of 0.25M. The reaction mix was neutralised by addition of HEPES free acid (Sigma, Dorset, UK, see table 2.1) to a final concentration of 0.06M. Clean-up was performed using CyScribe GFX™ purification columns (included in the kit) according to manufacturer's instructions. Purified cDNA was stored at –20°C.

2.12.1.4 Qiagen LabelStar™ RNA Labelling Kit

LabelStar™ complete kit was purchased from Qiagen (Qiagen, West Sussex, UK). The kit included reagents for generating labelled RNA and subsequent clean up of reaction products. Cy3/Cy5 labelled dUTP were purchased from Amersham Biosciences (Bucks., UK). Manufacturers protocol was followed throughout the process. Purified cDNA was stored at –20°C.

2.12.1.5 Micromax™ ASAP Labelling Kit

The Micromax™ ASAP Labelling Kit was supplied by Perkin Elmer™, MA, USA. Again, Cy3/Cy5 labelled dUTP were purchased from Amersham Biosciences, UK and the manufacturer's protocol was adhered to. Following the reaction cDNA was cleaned using an Oligotex™ mRNA spin column (Qiagen, West Sussex, UK) according to manufacturer's recommendations.

2.12.2 Fluorescent Labelling of Antisense RNA

The fluorescent labelling of aRNA, generated as described (section 2.8.6), was the same as that described for total RNA (section 2.12.1.3). However as the aRNA lacks a poly-A tail only random hexamers were used to prime the template.

2.13 BACTERIAL CULTURE MEDIA PREPARATION

The correct number of LB Agar/ Broth tablets (Sigma, Dorset, UK) were added to a specific volume of Milli-Q H_2O according to the manufacturers instructions. The tablets were dissolved by autoclaving in a a Prestige Medical 2100 benchtop autoclave (Prestige Medical, Lancs., UK). The solutions were allowed to cool to approximately 50°C and ampicillin antibiotic was added to a final concentration of 100µg/ml.

Agar plates were prepared by pouring approximately 40mls of the LB Agar into a Petri dish (Sterilin, Staffs., UK). Once solidified the plates were stored at 4°C. LB Broth was also stored at 4°C.

2.14 GENE CLONING

2.14.1 PREPARATION OF DNA PRIOR TO CLONING

Primers speficic to the gene of interest were designed and PCR amplification of the desired gene was carried out (section 2.9.2) using cDNA, generated from cortical total RNA (section 3.3.3) as template. Prior to cloning the PCR fragment was cleaned to remove components of the reaction mix (section 2.9.3).

2.14.2 LIGATION REACTION

Purified PCR generated DNA was cloned into the pGEM®-T vector system (Promega Ltd., Southampton, UK) by the utilisation of 3'-Thymidine (T) overhangs produced by EcoRV digestion of the pGEM®-5Zf(+) vector followed by the addition of a T to both 3' ends. The ability of Taq DNA polymerase to add a single adenosine (A) to the 3'-end of PCR products subsequently facilitates the cloning of the template into these sites. Sterile microfuge tubes containing the supplied pGEM®-T Vector (Promega Ltd., Southampton, UK) were briefly centrifuged in a desktop centrifuge to collect the contents at the bottom of the tube. The ligation reaction was set up by the addition of 1x T4 DNA Ligase Buffer, 50ng pGEM®-T vector and 3U T4 DNA Ligase into a sterile microfuge tube. PCR

product, 100–200ng, was added and the contents of the tube were mixed by pipetting, and incubated at 4°C for 16 hours.

2.14.3 TRANSFORMATION OF COMPETENT CELLS

Following ligation PCR amplicons and the ligated vector for the relevant genes, were transformed into E. coli DH5α competent cells. Aliquots of DH5α competent cells (Gibco BRL, Paisley, UK) were thawed on ice. The ligated PCR sample (3µl) was added to the sterile microfuge tubes containing the thawed 50µl competent cells. The solution was placed on ice for 30 minutes, then heat shocked at 37°C for exactly 20 seconds. The samples were placed on ice for a further 2 minutes, then 0.95mls LB-broth was added at room temperature. The samples were then incubated in a 37°C orbital incubator, set at 225rpm for 1 hour. LB-agar (amp+) was prepared as described previously (section 2.13), and 150µl of the transformed cells plated out with a sterile spreader. These plates were incubated overnight in a 37°C incubator.

2.14.4 GENERATION OF GLYCEROL STOCKS

Individual colonies, generated following overnight incubation (section 2.14.3) were transferred to separate wells of a Falcon 96 well culture plate (Fisher Scientific, Loughborough, UK) containing 150µl of LB-agar (amp+) (section 2.13). After 3 hours of incubation at 37°C (or when a randomly picked well reached an OD of 0.8) 50µl from each well were transferred to the corresponding well of a new 96 well plate. 50µl of sterile 50% glycerol (VWR™ International, Dorset, UK) was added to each well of the new plate. The plate was mixed and this glycerol stock plate was stored at -80°C.

2.14.5 PCR SCREEN OF TRANSFORMED CELLS

Culture plates, from which glycerol stocks had been made, were spun down at 3000rpm for 5 minutes using a Beckman Coulter Allegra®21R Centrifuge fitted with plate holding trunnions. The supernatant was removed by inverting the plate. The pelleted cells were redissolved in 50µl of Milli-Q H_2O, and lysed by boiling the plate at 100°C for 10 minutes. The plates were respun to pellet the cell debris. The DNA from the ruptured cells was now free in solution and 1µl was transferred to the respective well of a 96 well PCR plate (ABgene, Epsom, UK). PCR amplification was carried out as described (section 2.9.2) using M13 primers specific to the pGEM®-T vector (table 2.2).

The PCR products were visualised using gel electrophoresis (section 2.10.2). Positively transformed colonies were identified as the larger of the two sizes of fragments. The positions of the positive cells on the agarose gel corresponded to the position of the positive cells on the glycerol plates thereby transformed cells could be traced to a glycerol stock.

2.14.6 PREPARATION OF PLASMID DNA FROM TRANSFORMED CELLS

Sterile loops were used to sample positively transformed cells, from the respective well on the glycerol stock plate, as identified by PCR screening (section 2.14.5). Overnight cultures, of transformed cells, were grown at 37 °C in LB (section 2.13), containing the appropriate antibiotic (100 μg/ml ampicillin for pGEM®-T vector), and centrifuged to produce a cell pellet. Plasmid DNA was isolated from overnight cultures, using the Wizard® *plus* SV miniprep purification system (Promega Ltd., Southampton, UK), according to the manufacturer's protocol. The supernatant was removed and the cell pellet resuspended in cell resuspension solution (Table 2.1). The EDTA makes the bacterial outer membrane permeable, and contains Mg^{2+} which chelates DNases, preventing their action. The RNase A degrades any RNA present in the sample.

Cell lysis solution was added (Table 2.1), containing 0.2 M NaOH, which denatures proteins and disrupts chromosomes as well as cleaving RNA. Cell lysis solution contains 1 % SDS, a detergent, which lyses the cell membrane, allowing plasmids to escape. Alkaline protease solution was added prior to incubation of the total reaction solution, which breaks down bacterial proteins. Neutralization solution (Table 2.1) was added, containing guanidine hydrochloride and potassium acetate which form a precipitate with the proteins and chromosomal DNA. The plasmids remain in solution. Glacial acetic acid in this solution neutralizes the alkali from the cell lysis step. The reaction was mixed and centrifuged, resulting in a clear supernatant containing the plasmid. The precipitate formed contains proteins, chromosomal DNA and cell debris.

This clear supernatant was transferred into a Wizard® *plus* SV miniprep Spin Column, where, under high salt conditions, the nucleic acids bind to the resin in the spin column. Column wash (Table 2.1) was added and centrifuged through the column to wash away contaminating proteins and salts (this step was performed twice). The wash solution contains ethanol and a small amount of salt, so that mono-nucleotides are removed but plasmid DNA remains bound.

The wash solution was removed prior to the plasmid DNA being eluted from the column by addition of nuclease-free water (heated to 65 °C). This causes the DNA to be released from the resin, which was subsequently collected by centrifugation of the column

2.15 CREATION OF MICROARRAY SLIDES

Microarray slides were created by printing cDNA generated from bacterial clones by PCR, onto CMT-GAPS™ coated glass slides (Corning, NY, USA). Two types of microarray slides were used in this thesis; a small-scale custom microarray (Chapter 4) and a large-scale, NIA 15K microarray (Chapter 5).

2.15.1 CLONE ACQUISITION – SMALL SCALE MICROARRAY

The source of each clone used on the small-scale microarray is provided in appendix 4. The majority of clones were obtained from the NIA 15K clone set (section 3.4.1, National Institue of Ageing, Baltimore, MD, USA, http://lgsun.grc.nia.nih.gov/cDNA/15k.html), however a number of clones were obtained from the IMAGE clone set and by RT-PCR (Invitrogen, Paisley, UK). Therefore, appendix 4 provides the source of each clone and an ID that is used to identify clones within their repective clone set.

2.15.1.1 Generating Clones by RT-PCR

As can be seen in Appendix 1, a number of clones were generated by RT-PCR cloning procedures as described (section 2.14.1 – 2.14.5). These clones were stored as glycerol stocks at -80°C until required. The sequences of primers used to specifically amplify fragments of the required genes are provided in table 2.2.

2.15.1.2 Generating cDNA from IMAGE Consortium Clones

The IMAGE clone for the required gene was identified using the Unigene database (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene) at the NCBI. Not all clones contained the gene ie negatively transformed, therefore it was necessary to identify positively transformed cells. To this end individual colonies were grown, from the culture supplied, on LB Agar Amp + petri dishes (section 2.13). Individual colonies were picked and glycerol stocks were made (section 2.14.4). A PCR screen followed to identify positively transformed colonies in the glycerol stocks (section 2.14.5). The glycerol stocks were stored at -80°C until required.

2.15.1.3 Generating cDNA from NIA 15K Clone Set

The clones from the NIA 15K clone set were already PCR screened and were stored, as glycerol stocks at -80°C, until required.

2.15.2 RE-ARRAYING OF cDNA CLONES FOR CUSTOM CHIPS

For convenience of handling, fresh glycerol stocks of every cDNA clone, used on the custom microarray chip, were made in four new 96 well culture plates. Each cDNA clone was allocated a well on one of four 96 well culture plates (Fisher Scientific, Loughborough, UK). This meant that regardless of the source of the clone ie from the NIA 15K set, the IMAGE consortium or generation by RT-PCR, all clones used for the custom chip were together in an ordered format.

150µl of LB Broth (Amp+) was placed in the relevant well of one of the four new 96 well plates. Each well was innoculated with the respective clone. For RT-PCR generated clones or clones from the IMAGE consortium only positive cells from the PCR screen were used (sections 2.15.1.1 and 2.15.1.2). Individual clones required from the NIA 15K set were identified by using the 'H' number provided in Appendix 1 eg H3004A10 refers to 96 well plate 4, row A column 10. The four 96 well plates were incubated and glycerol stocks were made as described.

2.15.2.1 PCR Amplification of Re-Arrayed Clones

The re-arrayed clones in glycerol stocks (section 2.15.2) were used to generate further growth in four new plates. Following growth the cells were lysed to release DNA template (section 2.14.5). Taking care to maintain the layout of the plate PCR amplification of every template was carried out as described (section 2.14.5). The products were purified (section 2.9.4) and visualised by gel electrophoresis to ensure success of PCR prior to microarray printing (section 2.10.2).

2.15.2.2 Sequence Check of Re-Arrayed Clones

Purified PCR products (section 2.15.2.1) were transferred to a 384-well plate (ABgene, Epsom, UK) recording the location of each cDNA clone within the plate (see appendix 1 for list of plate coordinates). Five purified PCR products were chosen at random from the 384 well plate and sequenced as described (section 2.9.5) to ensure the correct clone was in the correct position within the plates prior to microarray formation. Plate coordinates

are provided in Appendix 1.

2.15.2.3 Preparation of cDNA Prior to microarray Printing

To each well was added an equal volume of DMSO (Fisher Scientific, Loughborough, UK) to denature the double-stranded DNA, preventing evaporation during printing and facilitating hybridisation with probe.

2.15.2.4 Printing of cDNA Products to Form Custom Microarray Slides

The custom microarrays were printed using an 8x 4 solid titanium pin head (Genomic Solutions, Cambridge, UK)with each pin depositing 300 pl of PCR product. Prior to use each pin was removed from the head and cleaned by wiping with tissue soaked with Milli-Q H_2O followed by tissue soaked with 95-100% ethanol. Care was taken not to touch any part of the pins.

The cDNA, generated from the re-arrayed clones (section 2.15.2.1 – 2.15.2.3) was spotted in a known format onto CMT-GAPS coated slides (Corning, NY, USA), using a Flexys™ High Density Arrayer (Genomic Solutions, Cambridge, UK). Prior to printing the user had to input the requirements of the printing run into the computer software that controlled the process. The user was prompted for answers to a number of questions such as number of slides to be printed, number of replicates of each spot per slide, number of 384 well plates etc. In addition, information about the use of control spots, such as blanks and landmarks, had to be specified. This information was used by the arrayer to print the cDNA clones onto the slide in an orderly manner and to generate a .csv file, which linked each spot on the slide to an identifier specified by the user.

Each cDNA clone was spotted four times on the custom microarray chip. The 32 pin head would pick up PCR product from the 384 well plate and in a predefined position deposit the cDNA clone. After placing each replicate in a known spot the pins of the head were cleaned using designated tools within the arrayer. Sequentially the pins were washed using a Sonicating Bath containing Milli-Q H_2O for 10 seconds, a Brushing Bath containing Milli-Q H_2O for 10 seconds and a Brushing Bath containing 75% ethanol for 10 seconds. The pins were then subjected to a halogen drying lamp for 10 seconds followed by air drying for 5 seconds. The head would then return to the 384 well plate and start to print the next batch of cDNAs. In total 2352 spots were present on each custom array, 384 x 4 known genes and 812 control spots. The array could be broken down into 12 x 4 sub arrays each consisting of 49 (7 x 7) cDNA spots (see figure 3.17).

2.15.3 PRINTING OF NIA 15K MICROARRAY SLIDES

2.15.3.1 Generation of cDNA from NIA 15K Clone Set

The NIA 15K clone set came supplied as sequence verified glycerol stocks. In addition cDNA was provided, generated by the NIA and supplied, ready to print, in two times volume of DMSO.

2.15.3.2 Generating Microarray Slides from the NIA 15K Clone Set

CMT-GAPS™ (Corning, NY, USA) coated microscope slides were used as the substrate for printing cDNA generated from every clone from the NIA 15K clone set. Printing was conducted using a Perkin Elmer® SpotArray 72 (Perkin Elmer®, MA, USA) printer. Prior to printing each one of the forty 384 well plates housing the cDNA derived from the clone set were given a unique ID. Information relating to the individual clones within each plate was imported into the print software as .txt files and became part of the SpotArray database where plate definitions are saved. Thus, the plate definitions contain detailed information about every clone contained in each plate. The plate definitions of the 40 plates were then combined to generate a plate set dataset which could be saved and re-opened for repeat print runs.

Before printing, printing protocols were defined using a wizard which guided the user through the various parameters to be used. Sequentially the following information was required:

Basic Information: *A name and description of the protocol*

The plate set to be used

Pin Selection: *The pin configuration is selected (in this case for 48 pins) and a graphical representation showing the position of the pins in the print head generated. The pins were installed in the printhead so that their layout matched the illustration.*

Number of prints: *The number of microarrays to be printed*

Array Layout: *Number of replicate samples per array (in this case 0).*

Spot Spacing - the distance in (μm) between the centre of each spot. (Spacing was specified as 1.5 times the diameter of each spot (120μm in this case).

Location of the array on the slide (centered).

Sub-array spacing, the distance between each sub-array (40μm)

Pin Motion: *Time taken to print slides - composed of time pins spend on slide (dwell time), approach velocity of pins to array, sample load dwell time (amount of product taken up by pins) thickness of array etc. All parameters serve to define the amount of cDNA deposited at each spot (default values used) and number of spots that can be printed from a single uptake.*

Washing and Drying: *Length and number of washes and length of drying time following washes.*

Environmental: *Humidity check - to identify conditions suitable for printing. Temperature control. Essentially a thermometer that prevents printing if temperature is outside the acceptable range.*

Following the creation of a print protocol split pins were placed in the printing head, in the correct configuration and cleaned a minimum of three times using the Wash and Dry Pins option on the print control software. Microarray slides and 384 well plates containing cDNA generated from the NIA clone set were placed in the printer in the required position. The printer then checked the positioning of slides and plates matched the specifications of the print protocol using an optical sensor and printing was initiated.

2.15.4 STABILISATION OF cDNA ON CMT-GAPS™ COATED SLIDES

Immobilisation of the cDNA onto the CMT-GAPS™ coated slides, for both the small-scale and NIA 15K microarrays, was achieved by means of baking and UV crosslinking. The DNA on the printed slides was initially immobilised by placing in a UV Stratalinker™ 2400 bench top transilluminator (Stratagene Ltd.) for five minutes. The microarray was then placed in a lightproof slide container and baked in an oven at 80°C for 2 hours. If not used immediately, the microarrays were stored desiccated at room temperature, protected from light.

2.15.5 QUALITY CONTROL OF MICROARRAY PRINTING

For each batch of 10 slides printed, one was chosen, at random, for quality control measures. SYBR® Green II nucleic acid gel stain (Molecular Probes, Cambridge Biosciences, UK) was diluted 1:10,000 in TE buffer (see table 2.1). The chosen microarray slide was immersed in this solution at room temperature for 2 – 3 minutes. The slide was rinsed in untreated TE buffer followed by Milli-Q H_2O. The slides were then scanned, using a GeneTAC™LS-IV scanner (Genomic Solutions, UK), at 600nm (section 2.17).

2.16 HYBRIDISATION

2.16.1 PROBE PRE-HYBRIDISATION TREATMENTS

Labelled probes were prepared as described (section 2.12.1.3). The probes required for each hybridisation were combined in sterile amber eppendorf tubes (Fisher, UK). The volume in each tube was reduced to approximately 3µl using an Eppendorf Concentrator 5301 (Jencons, UK) followed by addition of 1 x hybridisation buffer (50% formamide, 5 X SSC, 0.1% SDS) to make a final volume of 30µl. Poly dA^{30} (2µl of 20µg/µl, MWG, Germany) and mouse COT1-DNA (2µl of 20µg/µl, Invitrogen™ Life Technologies, UK were added to prevent non-specific hybridisation of probe to target. In cases where the GeneTAC™ Hybridisation Station (Genomic Solutions, Cambridge, UK, section 2.16.3.1) was used 110µl of hybridisation buffer was added to the dried pellet. The solution was heated at 95°C for 5 minutes to denature the probe and snap cooled on ice.

2.16.2 MICROARRAY SLIDE PRE-HYBRIDISATION BLOCKING

Prior to hybridisation it was found necessary to treat the printed microarray slide to prevent non-specific binding of probe to reactive amine groups in the Gamma Amino Propyl Silane coating. This was achieved by placing the microarray slides in 50ml Coplin jars, containing blocking solution (1% Bovine Serum Albumin (Fraction V) (Sigma-Aldrich, Poole, UK), 5x SSC and 0.1% SDS), pre-warmed to 42°C. Ensuring the slide was completely covered with blocking solution, the Coplin jars were incubated in a water bath, set at 42°C, for 45 minutes. The microarray was then dipped briefly (2 seconds each) five times in sterile, filtered Milli-Q H_2O. This was followed by a single 5-second dip in room temperature isopropanol followed by spin drying at room temperature at 2500 rpm in a Beckman Coulter Allegra™21R Centrifuge. Following reaction, it was imperative that the microarray be used for a hybridisation within 1 hour of blocking, since hybridisation efficiency decreases rapidly if left for any longer (Hedge *et al.*, 2000).

2.16.3 METHOD OF HYBRIDISATION

Hybridisation involved holding the microarray slide, overlaid with labelled probe, at 42°C for 16 hours. Two hybridisation methods were tested; manual and automated.

2.16.3.1 Manual Method of Hybridisation

To minimise dust specks on the microarray slide powder free gloves were used during microarrray slide handling. Suitably blocked microarray slides were treated with compressed air (KenAir Air Duster, Kenro, UK) followed by static charge neutralisation (Aldrich ZeroStat®, Sigma, UK) to remove specks of dust. The microarray slides was then placed, array side up, onto a dark surface and taking care not to introduce bubbles, the labelled probe was pipetted onto one end of the slides. Working quickly pre-dusted microscope slide cover slips (22 mm x 60 mm) were placed near the probe and lowered until the liquid wicked to the surface. The opposite end of the cover slips were gently lowered spreading the probe under the cover slip and covering the printed area of the microarray slides. During this stage it was critical not to introduce air bubbles. If formed, larger bubbles could be removed by gently pressing the cover slip. Immediately the covered microarrays were transferred to a light tight microscope slide chamber (Sigma, UK) in an oven set at 42°C. The chamber had previously been filled with Milli-Q H_2O just below the level of a raised platform. The microarray slides sat on the raised

platform so that they were in a constantly humid environment for the hybridisation of 16 hours. It was essential that the slides did not come into direct contact with the water.

Following hybridisation the cover slide was removed by dipping each slide in a coplin jar containing pre-warmed separating solution (1X SSC, 0.2% SDS see) so that the cover slip slid from the array. Non-specific contaminants bound to the arrays were removed by washing. Each slide was placed into a light tight coplin jar containing one of three pre-warmed wash solutions. The first wash used a high stringency buffer (2 X SSC) that removed all components of the hybridisation buffer. The coplin jar was shaken for 5 minutes (Stuart Rocking Platform, Jencons, UK) before repeating the wash. The next wash solution was a low stringency buffer (0.1 SSC, 0.1% SDS) designed to remove salt and loosely bound probe. Again the slides were shaken for 5 minutes and the wash repeated. Finally, low stringency buffer (0.1X SSC) was used to wash the slides twice with shaking.

All wash solutions were pre-warmed to 42°C and between washes microarray light exposure was kept to a minimum. After washing, the microarrays were then dipped briefly (2 seconds each) five times in sterile, filtered Milli-Q H_2O. This was followed by a single 5-second dip in room temperature isopropanol followed by spin drying at room temperature at 2500 rpm in a Beckman Coulter Allegra™21R Centrifuge. Prior to scanning the slides were stored at 4°C in light tight containers.

2.16.3.2 Automated Method of Hybridisation

Prior to use, each chamber cover of a GeneTAC™ Hybridisation Station (Genomic Solutions, Cambridge, UK) was removed and wiped with 0.1 X SSC (Sigma, UK) while each sealing rubber ring was dipped twice in boiling Milli-Q H_2O and dried. Suitably blocked microarray slides were dusted (as in section 2.16.3.1) and placed array side up on the bottom holding plate. The rubber sealing ring was placed back into the chamber cover which was then fixed back onto the hybridisation station. Each cover was lowered and tightened onto the microarray slide. Prior to introduction of the probe the microarray slides were held at 42°C. Suitably prepared probes were applied to the chamber by pipetting through a small, pluggable hole.

Following hybridisation the station was programmed to wash each microarray slide. Wash 1 solution (1 X SSC, 0.2% SDS) and wash 2 solution (0.1X SSC, 0.2% SDS) were fed into each chamber from a master bottle. Each microarray slide was washed twice

with each solution for 10minutes at 42°C. Following each wash the solution was removed to a waste bottle and fresh solution was fed in.

After washing, the microarrays were then dipped briefly (2 seconds each) five times in sterile, filtered Milli-Q H_2O. This was followed by a single 5-second dip in room temperature isopropanol followed by spin drying at room temperature at 2500 rpm in a Beckman Coulter Allegra™21R Centrifuge. Prior to scanning the slides were stored at 4°C in light tight containers.

2.17 SCANNING OF MICROARRAY SLIDES

All microarray slides, and agarose gels were scanned using a GeneTac™ LSIV scanner (Genomics Solutions®, Cambridge, UK). Prior to scanning microarray slides were treated with compressed air (KenAir Air Duster, Kenro, UK) followed by static charge neutralising (Aldrich ZeroStat®, Sigma, UK) to remove specks of dust. Ensuring correct orientation, slides and gels were loaded onto a 24 unit carousel in the scanner and the door closed. The user was required to input the ID of each scanned slide using GeneTac™ Software, which accompanied the scanner. Before slides were scanned parameters needed to be set that enabled the laser in the scanner to be tuned to the excitation wavelengths of the fluorophores under investigation, namely Cy3 (550nm, false colour green) and Cy5 (650nm, false colour red). Data was typically collected in two different channels, the Cy3 or the Cy5 channel but up to four channels could be specified. Initial, preview, scans were carried out on each microarray using a low resolution setting of 50μm. Using a low resolution ensured rapid scanning and an image was generated on the computer VDU of each channel in addition to a composite image of the two channels.

During preview scans Gain and Offset of the laser were adjusted using the GeneTac™ software so that the optimum image quality for the series of scans could be found. Gain values were adjusted so that the brightest spots on the array were just saturated (appear slightly white) while offset values were set to the minimum value that enhanced dim spots and kept background flourescence to a minimum. For final scanning the highest possible resolution was used (10μm) and using this setting the average scan time for a single channel was approximately 40 minutes. Each microarray image was saved as 16 bit single channel TIFF files.

2.18 SCANNED IMAGE ANALYSIS

Scanned TIFF file images were imported directly into Imagene™ (Biodiscovery, Inc. CA, USA) microarray analysis software, which located spots and quantified the amount of fluorescence at each spot. This process is and the use of Imagene is described in more detail in sections 3.7.2 – 3.7.3.

2.19 SOFTWARE USED IN DATA ANALYSIS

2.19.1 DATA PAIRING AND TRANSFORMATION

Quantified data from each channel of the scanned microarray images were paired and transformations were carried out on the resulting data using GeneSight® software suite (BioDiscovery, Inc.,CA). Use of this software and transformation parameters used prior to analysis of data are described in detail in sections 3.7.2 and 3.8.

2.19.2 STATISTICAL ANALYSIS OF MICROARRAY DATA

The Statistical Analysis of Microarray (SAM) was used to identify significantly differentially expressed genes in Chapter 5. SAM is available for free download from http://www-stat.stanford.edu/~tibs/SAM/ complete with an in-depth manual for use, tutorials and test datasets. The software works through Microsoft Excel as a simple plug-in and its operation is described thoroughly in section 5.6.4.3 – 5.6.4.4.

2.19.3 FUNCTIONAL ANALYSIS OF MICROARRAY DATA

The distribution of functional groups within significant gene lists was carried out using DAVID software. The significance of these functional groupings was carried out using EASE. Both DAVID and EASE can be accessed online at http://david.niaid.nih.gov/david/ease.htm .

2.20 REAL TIME QUANTITATIVE-PCR

PCR based quantitative amplifications were performed to assess the expression levels of specific gene sequences in experimental samples. The TaqMan® 5′-Exonuclease Probe quantitative PCR technique (qPCR) measures the accumulation of PCR products through a dual-labelled fluorogenic probe (TaqMan® probe), allowing the quantitation and expression profiles of genes to be determined (Holland *et al.*, 1991; Nasarabadi *et al.*,

1999). The TaqMan® probe technology was utilized in the ABI Prism® 7700 sequence detection system.

The exonuclease probe, non-extendable at the 3′ end, is dual-labelled, with the reporter dye 6-carboxyfluorescein (FAM) at the 5′ termini, and 6-carboxy-tetramethyl-rhodamine (TAMRA) at the 3′ termini as the universal quencher (Nasarabadi *et al.*, 1999). The spatial proximity of the reporter dye to the universal quencher quenches the fluorescent emission. Primers are designed to amplify a region of approximately 150 bp, and the probe is designed to hybridize within the target sequence. During amplification the hybridized probe is digested by the 5′ → 3′-exonuclease activity of the *Taq* DNA polymerase in the primer extension phase of amplification, thereby releasing the reporter dye, resulting in a relative increase in fluorescent signal (Nasarabadi *et al.*, 1999).

2.20.1 PRIMERS AND TAQMAN® PROBE DESIGN

Primers and TaqMan® probes were designed, for the genes to be quantified, using the Primer Express™ software package (PE Applied Biosystems). The primers were synthesised by MWG Biotech (Erbesberg, Germany). The TaqMan® probes were synthesised by Oswel (Southampton, UK), with the reporter dye FAM covalently linked to the 5′ ends and the quencher dye TAMRA at the 3′ ends, which were phosphorylated to prevent probe extension.

Oligonucleotide primers and TaqMan® probes were designed to amplify an amplicon of approximately 150 bp in length, with the TaqMan® probe hybridizing to the sequence. There are strict guide-lines to be followed when designing primers and probes:

i. Primer guidelines:

α) Tm of primers must be between 58 and 60°C.

β) The primers must contain between 20 and 80 % GC content.

χ) The length of the primers must be between 9 - 40 bases.

δ) There must be <2°C difference in the Tm of the two primers.

ε) Maximum of 2 G/C at the 3′ end.

ii. Probe guidelines:

α) Tm should be 10°C higher than the Tm of the Primers.

β) The probe should contain between 20 - 80 % GC content.

χ) The length of the probe should be between 9 - 40 bases.

δ) There should be no G on the 5′ end of the probe.

ε) There must be no more than 4 contiguous G's.

ϕ) The probe must not contain more G's than C's.

iii. Amplicon:

α) The amplicon should be between 50 - 150 bp in length.

β) The 3′ end of the primer should be as close to the probe as possible without overlapping.

In addition to these parameters TaqMan® probes were designed across intron / exon splice junctions (section 5.6.5.3). Table 2.3 provides the LocusLink ID of the genes assayed which enable the mRNA sequence and genomic clone information to be obtained from the LocusLink database. Table 2.3 also provides probe and primer sequences for qPCR assays.

2.20.2 GENERATION OF CALIBRATION STANDARDS

Calibration standards were created by cloning each gene required. This was achieved using cDNA, generated from the mouse cortex (section 3.3.3) by reverse transcription (section 2.8.7). PCR amplification of each gene was carried out and products were cloned into competent DH5α cells as described (section 2.14.1 – 2.14.4). Positively transformed cells were identified in glycerol stock plates, through a PCR screen, and plasmid DNA was extracted (section 2.14.6) and quantified. A standard dilution series of calibration standards were prepared by diluting each purified plasmid DNA to a known concentration, in the range 1 ng/μl to 1 fg/μl.

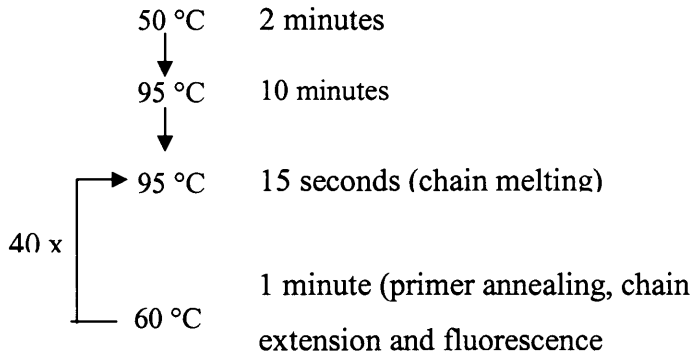
2.20.3 OPTIMIZATION OF TAQMAN® RT-PCR AMPLIFICATION

In order to generate the maximum information from the qPCR assay a number of variables needed optimisation prior to running experimental samples. Optimisation of two variables was necessary for this technique of TaqMan® probe real-time PCR quantitation in the ABI Prism® 7700 sequence detection system:

Oligonucleotide Primer concentrations, and

Probe concentration.

The amplification reactions consisted of the following prior to optimization: 2.5 µl of template DNA (plasmid cloned DNA preparations of 1 ng/µl concentration), 0.9 µM concentration of each primer (5′ and 3′ primers, synthesised by MWG Biotech, Germany), 200 nM probe, 0.2 mM dATP, dCTP and dGTP, 0.4 mM dUPT, 1 Unit of inhouse Taq polymerase buffered in 20 mM Tris-HCl (pH 8.3),50 mM KCl, 4 mM $MgCl_2$, 1 µM ROX The PCR reactions were performed in 25 µl volumes in PCR 96-well plates, using optical lids, and subjected to the following thermal cycling parameters using the ABI Prism® 7700 Sequence Detection System:



Product formation was monitored at the end of each extension step by measuring the fluorescence emitted from the TaqMan® probes. The number of cycles required to reach a critical threshold of fluorescence (Ct values) was calculated.

i. Oligonucleotide primer concentrations

The quantification reactions were performed as described above. Combinations of three primer concentrations were tested. The reaction with the lowest Ct value and best amplification rate was taken as the optimum reaction conditions.

ii. Probe concentration

Eight different probe concentrations were used in the probe optimization reactions (25, 50, 75, 100, 125, 150, 175, 200 nM) performed as described above, using the optimized primer concentrations determined previously. The reaction with the lowest Ct value and best amplification rate was taken as the optimum reaction conditions. Once the probe

concentration had been optimized the technique was applied to analysing expression in experimental samples.

2.20.3 SAMPLES ASSAYED IN QPCR ANALYSIS

Total RNA was extracted from the barrel cortex of mice subjected to three forms of vibrissal deprivation and used to generate cDNA (section 2.8.7) for use in qPCR analysis. The experimental samples are described in further detail in section 5.6.2. Experimental samples were assayed in triplicate, for each gene, in parallel with triplicate replicate calibration standards. The amplification reactions were performed using the optimised conditions specific to the gene under analysis.

2.20.4 QUANTITATIVE ANALYSIS OF GENE EXPRESSION LEVELS IN EXPERIMENTAL SAMPLES

Data analysis was carried out using specialised software within the ABI Prism® 7700 Sequence Detection System. From the calibration standards a standard curve was obtained and used to produce a regression line, by plotting the cycle number required to attain a threshold fluorescence pertaining to the mid logarithmic portion of the amplification against $\log_{10}$ [molecules of target gene]. Using the standard curve obtained (described above), a regression line was generated, extrapolating each sample over the standard range to quantify the expression levels of each gene in each sample.

2.21 TABLES

All routine laboratory solutions were prepared using ddH2O. Sterilisation was achieved by autoclaving where required. Heat sensitive components were passed through 0.22 μm Nucleopore™ filters and added separately following autoclaving. Where required β-mercaptoethanol and DTT were always added fresh to any solution. EDTA, TAE and TE were made according to the protocol of Sambrook *et al.*, (1989). A list of routinely employed solutions and reagents is shown in Table 2.1. Table 2.3 gives a list of primer sequences used in gene cloning and optimised $MgCl_2$ and primer annealing PCR conditions. Table 2.4 provides the LocusLink ID of all gene assayed using qPCR with primer and TaqMan® probe sequences. In table 2.4 is a description of the different primer concentrations used to optimise qPCR reactions.

Procedure	Solution	Components
Cytochrome oxidase stain of barrel cortex	**Phosphate buffer (0.1m)**	Milli-Q H2O, 0.1M dibasic sodium phosphate, 0.1M monobasic anhydrous sodium phosphate, 0.5M sodium chloride
	Fixative	Paraformaldehyde 4%w/v, Phosphate buffer 0.1M
	cytochrome oxidase stain	Phosphate buffer 0.1M, sucrose 4.4% w/v, cytochrome c 0.05% w/v, diaminobenzidine 0.05% w/v
	Subbing solution	1% w/v gelatin, 0.1%w/v chromium potassium sulphate, Milli-Q H2O
Precipitation / Concentration of RNA Samples	**Sodium Acetate pH4.2**	NaOAc 3M, Glacial acetic acid, Milli-Q H2O
DNAse Treatment of RNA Samples	**RQ1 DNAse 10X buffer**	400mM Tris-HCl pH8.0,100mM MgSO4, 10mM CaCl2
	RQ1 DNAse Stop solution	20mM EGTA pH8.0
Agarose gel electrophoresis	**Gel Electrophoresis Loading Dye**	50% (v/v) Glycerol, 1% (w/v) Bromophenol blue, 1x TAE (sterilised)
	10X TAE Buffer	40 mM Tris-acetate, 1 mM Na2-EDTA (pH 7.6)
RNA Extraction	**Tri Reagent™**	4M guanidium thiocyanite, 25mM sodium citrate pH7.0, 0.5% sarcosyl, 0.1M 2-mercaptoethanol, 0.2M NaOAc pH4.2, phenol, chloroform (1:0.1:1:0.2)
Removal of brain	**ACSF**	124mM NaCl, 2mM MgSO4, 2mM CaCl2, 5mM KCl, 1.25mM NaH2PO4, 26mM NaHCO3, and 10mM glucose
Reverse Transcription PCR	**5 x MMLV RT Reaction Buffer**	50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl2, 10 mM DTT
Wizard® *plus* SV miniprep purification system	**Cell resuspension solution**	50 mM Tris (pH 7.5), 10 mM EDTA, 100 µg/ml RNase A
	Cell lysis solution	0.2 M NaOH, 1% SDS
	Neutralization solution	4.09 M Guanidine Hydrochloride, 0.759 M Potassium Acetate, 2.12 M Glacial Acetic Acid, (adjusted to pH 4.2)
	Column wash	60 mM Potassium Acetate, 10 mM Tris-HCl (pH 7.5), 60 % Ethanol
Real-Time Quantitative PCR	**10 x QPCR buffer**	200 mM Tris-HCl (pH 8.3), 500 mM KCl, 40 mM MgCl2, 10 µM ROX

Table 2.1 List of routinely employed solutions, reagents and media

GENE	T_M	$MgCl_2$ conc	GI accession number	Forward primer seq.	Reverse primer seq.
Amph	58	2.0	11560001	TTGAGGCCAAGTTT CACAAGG	CCTTTGTTGGGGTG ACTTTAG
Dynamin	58	2.0	6681206	CGAGGATCAGGAAT TGTCAC	GTGTGACGGCAAG AATGAGG
NGFbeta	58	2.5	7305312	ACCCGCAACATCAC TGTGGA	CCAGTGTTTGGAGT CGATGC
Pcdh12	58	2.0	8393915	TCTGGTGTCATTTG ACGTGC	CCTTGACCACGCTG ATTCC
Pcdh13	58	2.5	7657444	GGTGCAAGCTACAG ATAACG	GAGTTATCGTTCGT GTCTGC
Ras	58	2.5	7242161	TGGTGATTGATGGT GAGACC	AAAAGGCATCCTCC ACACC
Mtap2	60	1.5	6678943	AGGTCTCAGTTCAG TGCCAG	TCAGGCTTTGTCAC CTTCTC
GluR1	58	1.5	34328127	GCATCGGGTACCAC TACATC	TGTACTTTGGCCTC TTCCAG
GluR2	58	1.5	7305114	TGACTGCGAAAGGG ATAAAG	GTGTGTGCTCCAGG GTATTC
GluR3	58	1.5	8393312	AGCTCTGAAAATGG TGCAAG	CTACGATGGTCCTG TTCTCG
GluR4	58	1.5	9790090	CAGGGAATTGACAT GGAGAG	CGATAGCTGCTGTG TCATTG
NMDA R1	58	1.5	6680094	GTGGAATTCAATGA GGATGG	CATACACGAAGGG TTCTTGG
NMDA R2A	58	1.5	6680096	CATTGTCACCTTGG AGGAAG	CATTGGTCACCAGG TAGAGG
NMDA R2B	58	1.5	6680098	GTCCGAACACAGTT TCATCC	TCTGCAGGGACTTG TCTTTC
NMDA R2D	58	1.5	6680100	TTCTGGTAAACCCG TCACTG	AACCGAGTCTCTGA TGCAAG
CREB	58	1.5	19745143	GCTAACAATGGTAC GGATGG	AGGCTGTGTAGGA AGTGCTG
ACTIN	58	1.5	49865	TGCGTGACATCAAA GAGAAG	TACGGATGTCAACG TCACAC
GAPDH	58	1.5	6679936	CATCTTGGGCTACA CTGAGG	TTCAAGAGAGTAG GGAGGGC
M13	58	1.5	-	GTT TTC CCA GTC ACG ACG TT	CAG GAA ACA GCT ATG ACC ATG
IMAGE M13	60	1.5	-	TGT AAA ACG ACG GCC AGT	CAG CAA AGA GCT ATG ACC
OLIGO dT	-	-	-	TTTTTTTTTTTTTTT TTTT	
Src	60	2.5	6678128	TCTGCTGGACTTTCT CAAGG	CGTACAGAGCAGCT TCAGG

Table 2.2 Primer Sequences, Annealing Temperatures and $MgCl_2$ concentrations.

Table 2.3 Primer and TaqMan® probe Sequences for genes assayed by Real Time Quantitative PCR

Gene Name	Forward Primer	Reverse Primer	TaqMan Probe	LocusLink ID
Tgfb1i4	TCCACAGTCACCCATTTCATTG	TTTCTTCAAGCCAAAAGCTTATGAG	TCTACGAAAGAGACGTGAGACTCACACGCTG	21807
Lamr1	TGGAGCAGTACATCTACAAAAGGAA	AGCTGCGAGCAACAGCTTCT	AGTGACGGTATCTACATCATAAACCTGAAGAGGACC	16785
Tyro3	TGCTCCCTCTCCCTCTGTTTT	GTTGCGGGCTTCACAAGAA	ATGTGACAGGAGTGACCCAGCGCAC	22174
Gnb2-rs1	CAGCCCCAACCGCTACTG	TTCTTGCTTCAATTCATCTACAATGA	CCAGCATCAAGATCTGGGACTTGGAG	14694
Bsg	GGCAGAAGCGAGATCAATGTG	TCGGACTTGCAGACCAGTTTC	CCACCCAGGATCAAGGTCGGAAAGAAAT	12215
PPIA	TTCCTCCTTTCACAGAATTATTCCA	TGCCGCCAGTGCCATTA	ATTCATGTGCCAGGGTGGTGACTTTACAC	319427
Vps 4a	GATAACAGCGTGGTGTCCTTGTC	GCATACAGCTTCCCGTCTTTTC	CAGCCCAAGATGGATGAACTGCA	116733

CHAPTER THREE

USE OF MICROARRAY TECHNOLOGY: PROTOCOLS AND ANALYSIS

3.1 INTRODUCTION

Microarray technology lends itself particularly well to the study of plasticity in the mouse barrel cortex due to its high throughput nature and sensitivity. Generating data from a microarray experiment is a multi-step process. Errors introduced during these steps are additive and cannot be removed without repeating the experiment. The aim of this chapter is to optimise each step in the process in an effort to reduce potential sources of error and generate data of the highest possible quality.

Figure 3.1 provides an overview of the microarray procedure. RNA from two sources whose gene expression profiles are to be compared are extracted and labelled with distinct fluorophores (figure 3.1a). cDNA clones of known gene transcripts are spotted onto amino-silane coated glass slides (microarray or target). The printing is carried out using robotics under the control of specialised software which allows the user to identify the gene ID of each spot from it's location on the microarray (figure 3.1b). The labelled fluorphores (probes) are hybridised to the microarray slide and the slide is washed (figure 3.1c). The microarray is scanned using a tuneable laser and an image generated (figure 3.1d).

Using dedicated software packages the amount of fluorescence across the microarray slide at each spotted gene is quantified for both the fluorophores. Bioinformatic and statistical software packages transform this raw data allowing the gene expression profiles of the two RNA samples (figure 3.1 e to f) to be compared and contrasted.

3.2 PROBE PREPARATION

In the context of a microarray experiment the term probe refers to fluorescently labelled mRNA. In two colour cDNA microarray experiments two mRNA samples are labelled with different fluorophores allowing them to be distinguished when hybridised to the same microarray spot. This section aims to generate working protocols for the production of probe of a suitably high quality and yield for microarray hybridisations.

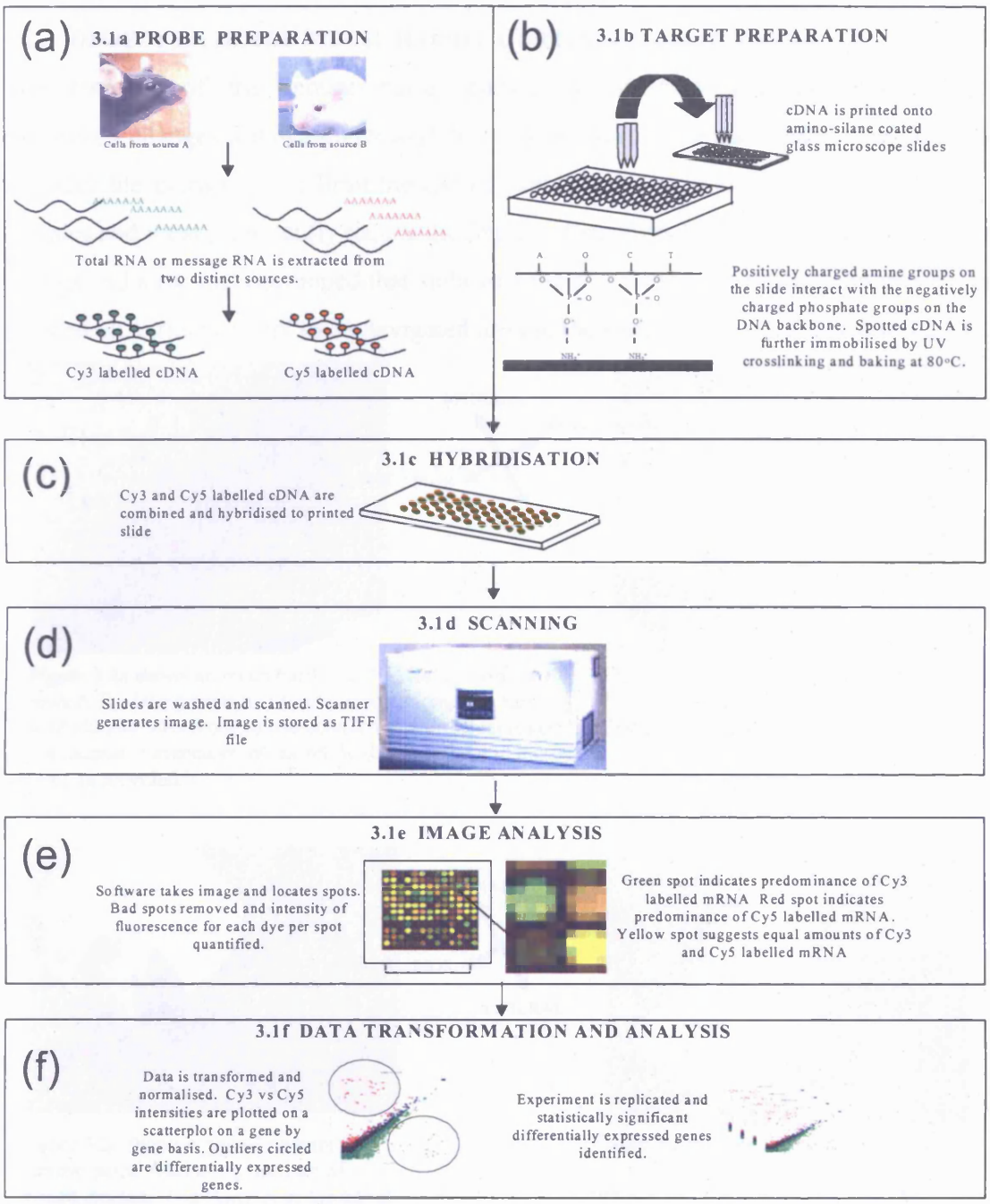


Figure 3.1 An overview of the microarray process.

(a) Probe production. mRNA from two experimental conditions whose gene expression profiles are to be compared are fluorescently labelled with different fluorophores.

(b) A cDNA microarray (target) is created. PCR fragments are spotted onto a glass microscope slide using specialised robotics.

(c) The probe and target are combined (hybridised) at 42°C.

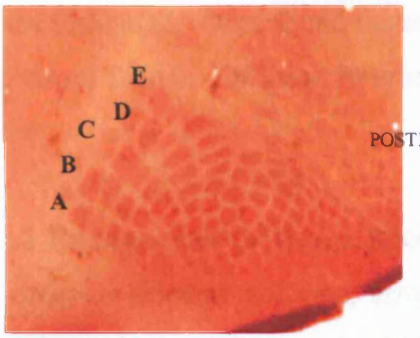
(d) The microarray is scanned at wavelengths appropriate for the fluorphores.

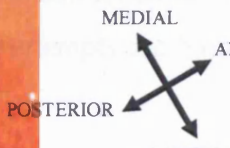
(e) The amount of fluorescence at each cDNA spot is quantified.

(f) The data is transformed and the experiment replicated to identify genes that are significantly differentially expressed between the two experimental conditions.

3.2.1 DISSECTION OF THE MOUSE BARREL CORTEX

The removal of the mouse barrel cortex is the first step in studying gene expression changes following vibrissal deprivation. A method was needed that facilitated reproducible extractions to limit the use of experimental animals, time spent processing samples and most importantly reduce the impact of sampling error on the data generated. To this end a rig was developed that stabilised the brain during the extraction process and allowed an extraction tool to be navigated around the surface of the brain (section 2.6).





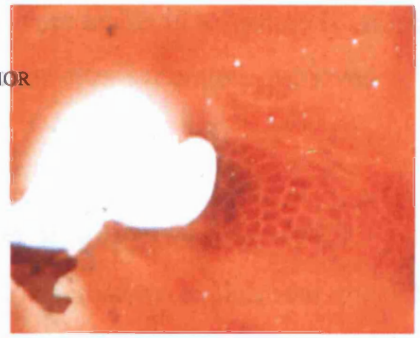
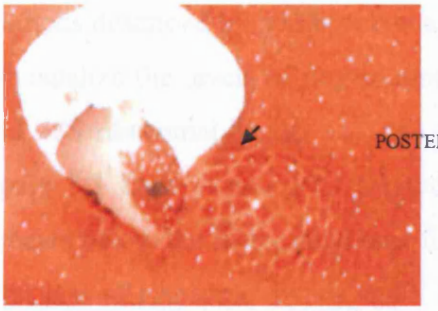


Figure 3.2a shows an intact barrel field. Clearly visible are rows A-E of the barrels that form the posteriomedial barrel subfield that form a one to one connection with whiskers on the muzzle. Orientation of barrel field with respect to mouse brain is provided.

Figure 3.2b shows many barrels remaining following an attempted extraction using first co-ordinates.



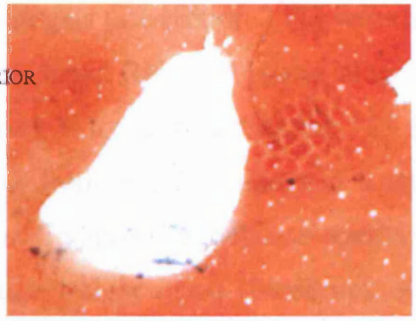


Figure 3.2c shows a typical attempt to remove barrel field using third set of measurements. The majority of the barrel field has been removed though some barrels of E-row remain (arrow)

Figure 3.2d shows a perfect removal of barrel field no barrels of PMBSF remain The co-ordinates used here were used for all further extractions (see section 2.4)

Figures 3.2 a-d show cytochrome oxidase staining of the mouse barrel cortex following extraction attempts. Knowing the orientation of the barrel field and visually analysing sections using a calibrated microscope lens, allowed precise co-ordinates to be calculated that enabled accurate sampling of the barrel cortex.

Initially, the location on the surface of the brain of a blue spot (corresponding to expression of a lacZ reporter gene) provided the co-ordinates of the D1 barrel from a landmark common to all mouse brains (Barth et al 2000, see figures 2.3 - 2.4 and section

2.6). Using these distances as a guide trials were carried out, using the rig to remove the barrel cortex (section 2.6). Animals were tested in batches and in total 4 co-ordinates were trialled until the optimum distance to move the borer from the landmark was found. Success of barrel cortex removal was assessed by the cytochrome oxidase staining of sectioned cortical tissue followed by visual inspection of the area sampled (section 2.7 and figure 3.2). Although the initial 3 co-ordinates failed to accurately remove the barrel cortex the area of extraction was similar in all animals used illustrating the reproducibility of the technique. Once found the correct co-ordinates were tested in a further 10 animals (20 barrel fields). Of these 20 attempts the barrel cortex was accurately removed in 15 cases.

3.2.2 EXTRACTION OF TOTAL RNA FROM THE MOUSE BARREL CORTEX

In order to generate accurate, error free data from a microarray experiment the RNA samples under investigation should be as intact and free from contaminants as possible. This section describes how working procedures were optimised for the extraction of total RNA from the mouse barrel cortex, using the tissue sections removed in section 3.2.1. Following the extraction process RNA integrity and purity were assessed using the methods described by Sambrook et al., (1989). Agarose gels (see section 2.10) were used to visualize the levels of degradation of each sample. Intact RNA has sharp, clear 18S and 28S ribosomal bands, with the 28S band appearing twice as bright as the 18S (see figure 3.3 Lane A and B). Degraded RNA deviates from this 2:1 ratio and a smear appears below the lower band (see figure 3.3 Lane C). Complete degradation results in a total loss of ribosomal bands and a large low molecular weight smear (Sambrook et al., 1989).

Initial attempts to extract total RNA from cortical tissue resulted in heavily degraded RNA samples (see figure 3.3 Lane C). Degradation was caused by RNAses, a family of enzymes responsible for the breakdown of RNA in nucleic acid metabolism (Levy, 1975). Found in prokaryotes and most eukaryotic cell types it was essential that RNA degradation was prevented, or at least detected, before continuing the process of probe generation. To this end techniques were established that minimised the chances of sample contamination with RNAses, as described below.

The post-mortem dissection of the mouse barrel cortex was carried out as quickly as possible to limit the potential exposure of tissue to RNAses and also to prevent any

changes in gene expression occurring within the dissected cells following the trauma of dissection.

Dissected tissue was immediately placed into TRI®Reagent (Sigma, UK), a solution which provided a rapid, cost efficient, one-step isolation of RNA (Chomczynski and Sacchi, 1987) (section 2.8.2). The TRI®Reagent contained guanidine thiocyanate, a strong denaturing agent which immediately denatured RNAses rendering them inactive. Where possible equipment was autoclaved prior to use and all surfaces were wiped down with a mild alkali solution. Using "RNA only" equipment in an area of the lab dedicated to RNA work minimised possible environmental contamination. All reagents were prepared using DEPC treated water (section 2.8). Any bought in products that came into contact with the RNA samples were certified as being RNAse free eg DNAse (Promega, UK). Care was taken to ensure that gel electrophoresis equipment was free of RNAses by soaking it in a solution of SDS (section 2.10.1).

The RNA samples in lanes A and B of figure 3.3 show the effect of these methods. Samples showing any degradation were disposed of.

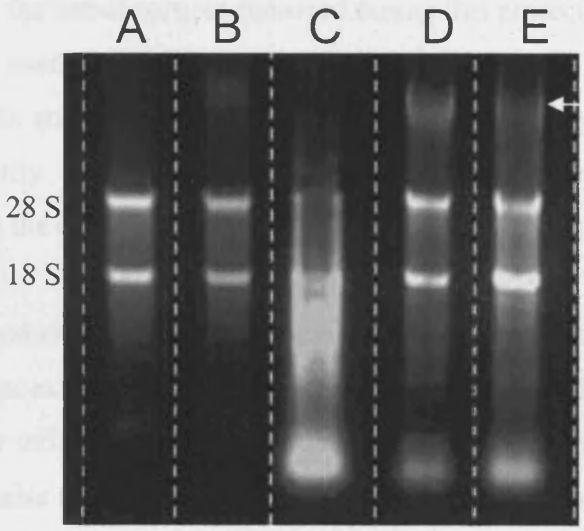


Figure 3.3 An agarose gel image of RNA samples removed from the area of the mouse barrel cortex.
Lanes A and B show intact RNA samples, the 28S ribosomal band being twice as bright as the 18S band. The sample in lane C is extremely degraded. 28S ribosomal band has almost disappeared and smearing appears below the 18S band.
Lanes D and E show presence of contaminating DNA (marked with an arrow) in the RNA sample.

The TRI®Reagent used in the extraction of RNA also had the capability to extract DNA and protein. Significant DNA contamination was observed in many RNA samples following gel electrophoresis (figure 3.3 Lane D and E). Contaminating DNA would

compete with labelled RNA during hybridisations and its removal was essential. In order to completely remove the DNA it was necessary to incubate each RNA sample with RNAse free DNAse (Promega, UK, section 2.8.4) followed by a sample clean up using Qiagen RNeasy® columns (section 2.8.5, figure 3.3, lanes A and B).

To measure the purity of each RNA sample the absorbance of UV light, of 260nm and 280nm wavelength, by each sample was determined using a spectrophotometer (see section 2.11). The optical density (OD) ratio OD260/OD280 for each sample was determined. Pure nucleic acid samples have a ratio between 1.8 and 2.1 with contaminating protein or chemicals carried over from the extraction protocol (eg phenol) significantly lowering the ratio (Sambrook et al 1989). Any RNA samples that fell out of this range were excluded from further work however, a Qiagen Rneasy® column clean-up was usually sufficient to bring the ratio of a sample into the acceptable range.

The OD 260nm was used to quantify the amount of RNA in each sample. An OD of 1 corresponds to 40μg/ml of RNA (Sambrook et al 1989). The spectrophotometer therefore measured the OD and converted it into a concentration. The average weight of total RNA obtained from the barrel cortices removed during this project was 2.4μg.

The quality control procedures of gel electrophoresis and measurement of the OD260/OD280, ensured that only RNA samples meeting the criteria of low degradation and high purity were used for analysis and prevented low quality RNA from compromising the experimental process.

3.2.3 CREATION OF AN RNA POOL FOR MICROARRAY OPTIMISATION

Total RNA was extracted from all tissue samples removed during the optimisation of the removal of the mouse barrel cortex (section 3.2.1). This not only provided samples with which to optimise the total RNA extraction protocol, but also allowed the formation of a total RNA pool, which was used throughout the project for various protocol optimisations. Only total RNA samples which passed the quality control procedures described were added to the pool (section 3.2.2).

3.2.4 GENERATION OF aRNA

One drawback of microarray analysis is the requirement of relatively large amounts of mRNA (approximately 1μg) or total RNA (approximately 10μg) in each hybridisation. The generation of anti-sense RNA (aRNA) is a well established technique that enriches the mRNA component of total RNA. Amplification is reported to produce a 10^6- fold

increase in the amount of mRNA (Van Gelder et al., 1990, Eberwine et al., 1992) and has been used to generate sufficient amounts of mRNA for the microarray analysis of single cells from within the CA1 region of the hippocampus (Kamme et al., 2003). Due to the limiting amounts of total RNA extracted from single mouse barrel cortices (section 3.3.2). RNA amplification was required for this project.

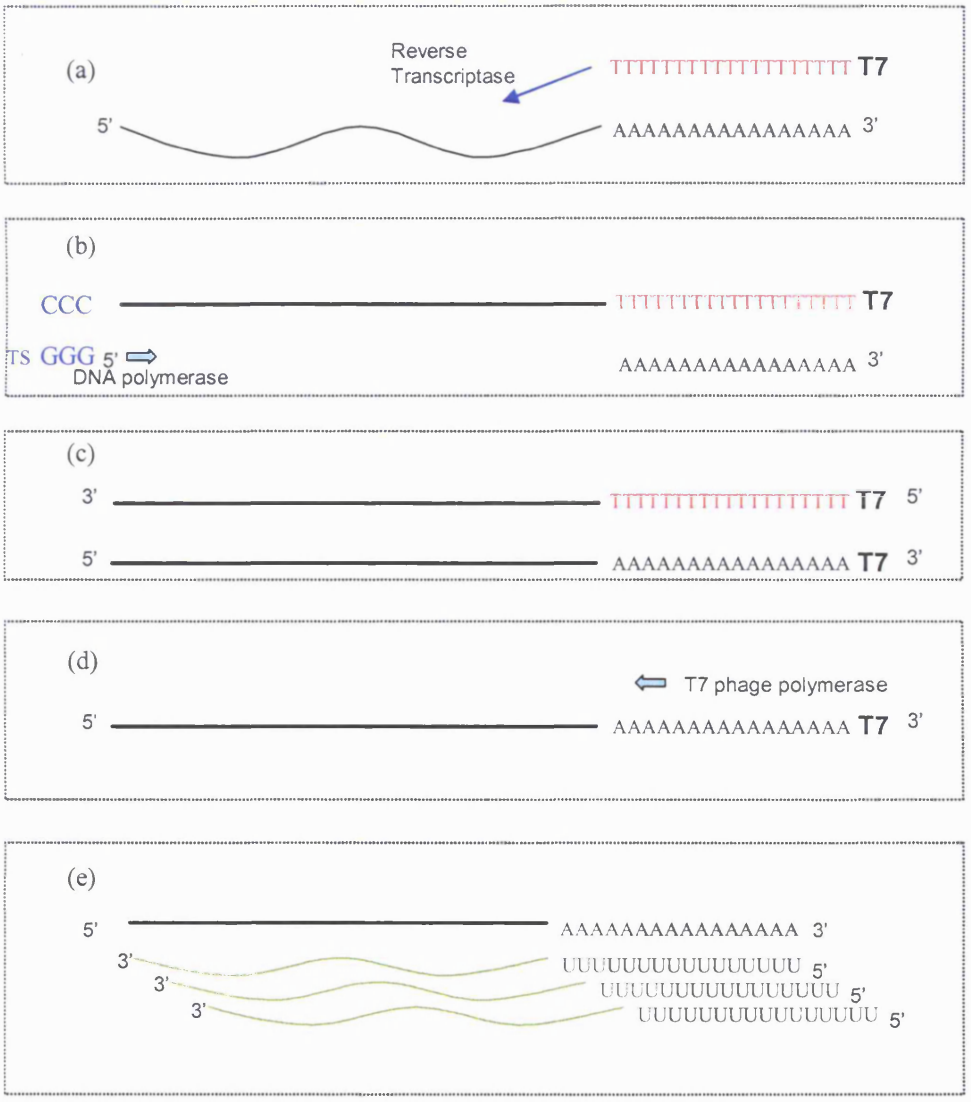


Figure 3.4 gives an overview of the generation of anti-sense (aRNA) using ™ aRNA kit from Ambion®.

(a) Total, or mRNA is primed with Oligo(dT) containing T7 promoter sequence. First strand cDNA is generated by reverse transcriptase.

(b) and (c) RNA is degraded and cDNA is primed using template switch (TS) primer complementary to the CCC overhang generated by reverse transcriptase. DNA polymerase generates double stranded DNA (dsDNA).

(d) and (e) T7 phage polymerase generates multiple copies of aRNA during in-vitro-transcription.

RNA amplification was carried out using a MessageAmp™ aRNA kit from Ambion®, Huntingdon, UK. An overview of the amplification reaction, based on the technique described in Van Gelder et al., 1990, is given in figure 3.4. The technique involves generating cDNA from total RNA, using oligo(dT) with a T7 promoter sequence as a primer for the enzyme reverse transcriptase (figure 3.4a). The formation of double stranded DNA (dsDNA) from the cDNA provides a template for a T7 phage poymerase to generate aRNA in an in-vitro transcription reaction (figure 3.4c and d). Aliqouts (1µg) of pooled total RNA (section 3.2.3) were amplified according to the manufacturer's protocols (section 2.8.6). Throughout the reaction care was taken to prevent exposure to RNAses (section 2.8.1).

The integrity, purity and concentration of the aRNA generated was determined as described for total RNA (section 3.2.2). The aRNA generated was run on an agarose gel next to a sample of the pooled total RNA (figure 3.5). The aRNA appears as a smear next to the total RNA with an average size less than 1.9kb (the size of the 18S ribosomal band, http://www.ambion.com/techlib/). The yield of aRNA was calculated to be 49µg. Given that only 1-3% of total RNA in mammalian cells is mRNA (Alberts et al., 2002) the reaction has resulted in an approximately 1500 fold increase of the transcript population. The OD260/OD280 ratio of the aRNA was 1.91 suggesting a pure sample free of contaminants.

3.2.5 FLUORESCENT LABELLING OF TOTAL RNA

Microarray technology allows the investigator to compare and contrast the abundance of specific gene transcripts contained in two separate RNA samples. The RNA samples are labelled with different fluorophores and applied to a glass slide onto which have been printed cDNA spots generated from known gene transcripts. Hybridisation occurs when transcripts within the RNA samples find a spot with complementary sequence. As both fluorophores have different excitation and emission spectra, transcripts that are present in both samples and hybridise to the same spot can be assigned to one or the other of the RNA samples. If the expression of a particular transcript is in equal abundance in both RNA samples then equal amounts of fluorescence will be generated at the spot. appear as a difference in the amount of fluorescence. Therefore, efficient labelling of RNA samples is a critical step in the microarray process. The labelling process should incorporate the different fluorophores equally and ensure all transcripts within an RNA sample are labelled (Schena et al, 1995, De Risi et al 1996).

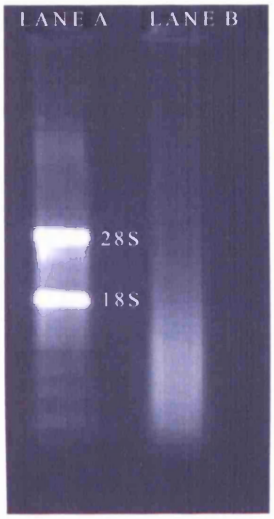


Figure 3.5 An agarose gel of aRNA (lane B) generated from pooled total RNA (lane A) using MessageAmp™ aRNA kit Ambion®, UK).

The two most commonly used in microarray work are Cy3 and Cy5, members of the cyanine family of fluorescent dyes (Ernst et al., 1989). Differences in the number of carbon atoms (marked with a black spot figure 3.6) in the bridge joining the two indocyanine rings (Mujumdar SR et al., 1996) means Cy3 and Cy5 molecules are excited at 532nm and 633nm respectively and following excitation show maximum light emission at wavelengths of 570nm (Cy3) and 670nm (Cy5) (Mujumdar RB et al., 1993).

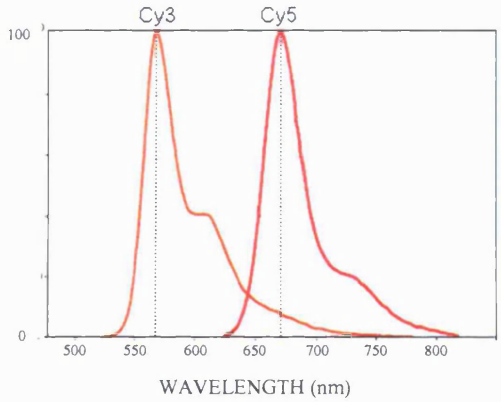


Figure 3.6 the emission spectra of Cy3 and Cy5 fluorescent dyes

Focusing a beam of light, at these excitatory wavelengths, at an array spot to which Cy3 and Cy5 labelled RNA is hybridised, generates fluorescence which can be assigned to either one or other of the RNA samples. Cy3 and Cy5 are used in the majority of microarray experiments as their emission spectra have negligible overlap and are therefore easily discernible with minimal cross-talk. The two fluorophores are also highly fluorescent when excited, are resistant to temperature and pH fluctuations and are relatively photostable (although care must be taken to minimise exposure) (Mujumdar RB

et al., 1993). Cy3 and Cy5 dyes were first coupled to nucleotides by Yu et al to produce labelled probes for fluorescence in-situ hybridization (Yu et al., 1994).

Protocols for the generation of Cy3/Cy5 labelled probe prior to microarray hybridization generally use the reverse transcription of total RNA or messenger RNA to form cDNA (see figure 3.7).

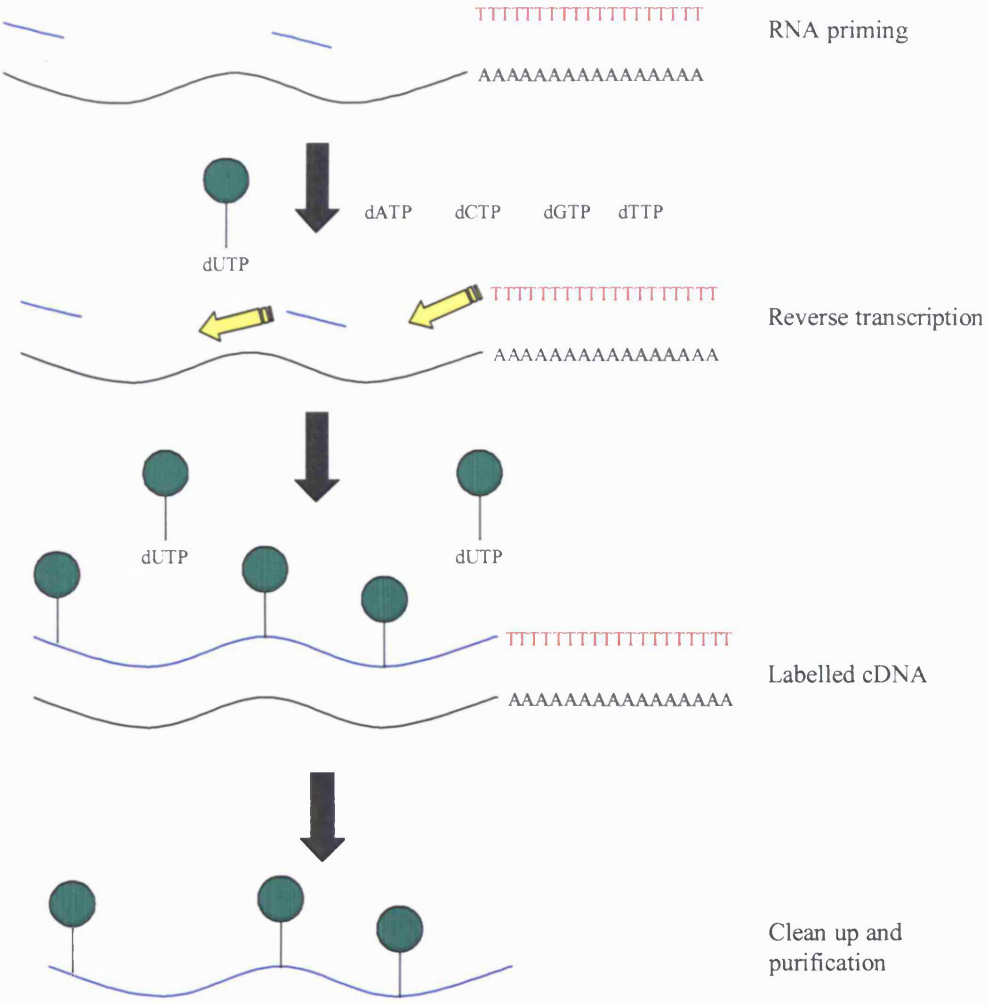


Figure 3.7 Direct labelling of RNA to form Cy labelled cDNA.

Oligo dT (red) and random hexamers (blue) are used to prime mRNA (black). The enzyme reverse transcriptase (yellow) adds complementary bases to fill gaps between primers. If a cyanine dye (green circle) is attached to a dUTP then this modified nucleotide will become incorporated during the reaction to form fluorescently labelled cDNA. A final clean up degrades RNA and removes unincorporated nucleotides.

Using the poly A tail, of messenger RNA as a priming site, means that theoretically all mRNA transcripts in a total RNA sample will be labelled. During the labelling process Cy labelled dNTPs are incorporated into the cDNA strand at random sites throughout the molecule. The dNTP mix must be balanced so that over-labelling and therefore fluorescent quenching does not occur.

The relatively high cost of Cyanine labelled dNTPs and the small amounts of RNA obtained from the barrel cortex meant that the labelling reaction needed to maximise the amount of labelled probe generated. A comparison was carried out between three commercially available labelling kits and a modified reverse transcription reaction to find the optimum method for production of probe. Reactions were tested using 10µg aliquots of pooled total RNA extracted from the cortex of mice (section 3.2.3). Protocols are described in section 2.12.1. Following the reaction the volumes of the generated products were equalised with H_2O. Equal amounts of products were resolved by gel electrophoresis and the incorporation of fluorophore was assessed by scanning the gel using a GeneTAC™LS-IV scanner (figure 3.8, section 2.17).

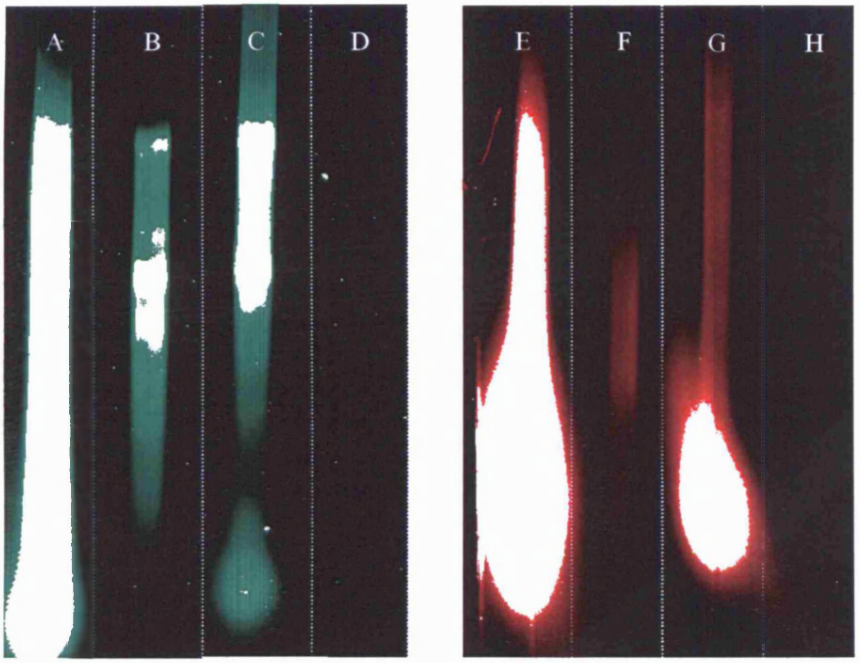


Figure 3.8 Scanned Images agarose gel loaded with 1ul of first strand labelling reaction products generated by various kits.

Lanes a-d Cy3 labelled cDNA

Lanes e-h Cy5 labelled cDNA

Lanes a and e Amersham Biosciences CyScribe™first strand labelling kit.

Lanes b and f Qiagen LabelStar™ Kit

Lanes c and g Modified Reverse Transcription kit

Lanes d and h Micromax™ ASAP Labelling Kit

The scanner assigns a false colour to each fluorophore in order to distinguish between Cy3 (green) and Cy5 (red) labelled product. This method of analysis is qualitative such that the amount of red or green smearing in each lane corresponds to the amount of labelled product. White smears are the result of the scanner reaching the limits of its dynamic range and is indicative of an excess of product. Thus it can be seen, quite clearly, that the CyScribe™ kit from Amersham Bioscience, UK forms the most labelled probe in both the Cy3 and Cy5 labelling reactions. Interestingly, the modified reverse transcription reaction, which utilised standard MMLV reverse transcriptase (Promega, UK) seemed to form more probe than either of the more expensive Micromax™ ASAP Labelling Kit or the Qiagen LabelStar™ RNA Labelling Kit. In all further studies the Amersham Bioscience CyScribe™ kit was used to form fluorescently labelled probe.

3.2.6 FLUORESCENT LABELLING OF AMPLIFIED RNA

The fluorescent labelling of aRNA with Cy3 or Cy5 fluorophores is essentially the same direct labelling reaction described for total RNA (section 3.2.5), the only differences being that random hexamers solely are used as primers and 1μg of aRNA is used in each labelling reaction (section 2.12.2). Oligo (dT) is not used as the aRNA is anti-sense and therefore lacks a poly-A tail.

aRNA generated from pooled total RNA (section 3.2.3) was fluorescently labelled with Cy3 and the probe generated compared to the probe formed directly from pooled total RNA (section 3.2.5). Both probes were run on an agarose gel and scanned as described (section 2.12.2). The image generated is shown in figure 3.9. It is clear from figure 3.9 that the aRNA has successfully incorporated the Cy3 fluorophore and generated probe that has a larger average molecular weight than probe from total RNA.

3.3 TARGET PREPARATION

In the mid 1990's the Brown lab developed a robotic system which printed cDNA clones (targets) of known genes onto glass microscope slides (Schena et al, 1995). The microarray field has evolved and moved on from in-house built printers to state of the art commercially available machines. Cardiff University's Molecular Biology Support Unit offered a core facility dedicated to the manufacture of cDNA arrays. The unit has two commercial printers, each of which prints cDNA from 384 well plates onto glass slides. Both printers offer different advantages to the user. Small scale (custom) microarrays can

be printed alongside high density microarrays containing (conceivably) the entire transcriptome of an organism. This section describes how protocols were optimised for generating high quality cDNA prior to printing a microarray chip.

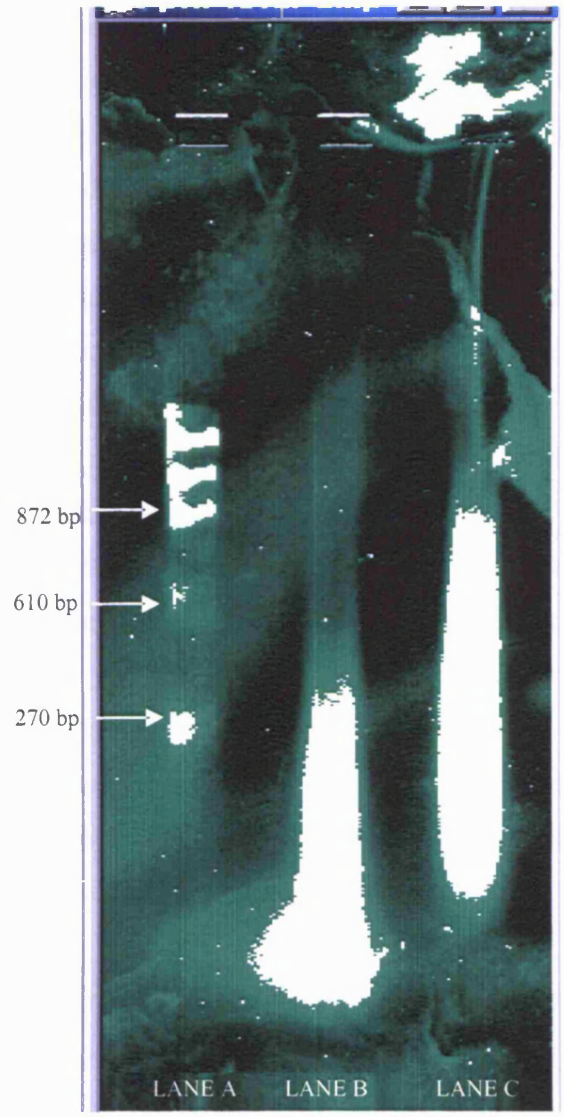


Figure 3.9 Agarose gel of fluorescently labelled probe formed from total RNA (lane B) and aRNA (lane C). Lane A is fluorescently labelled ØX174 size marker. Sizes shown in base pairs.

3.3.1 CLONE ACQUISITION

Before microarray production all transcripts (targets) under investigation are required in an ordered format that facilitates robotically controlled printing. The targets should be as free from contaminants as possible and in sufficient amounts to allow for multiple slide

printing runs. Two types of microarray chip were used in this project, a high density microarray (referred to as the NIA 15Kclone set array and a custom array).

The high density microarrays (section 5) were generated using the National Institute of Ageing's (Baltimore, MD, USA.) 15K mouse clone set provided by Cardiff University School of Biosciences. Their clone set was formed from 52374 expressed sequence tags (EST) that are primarily derived from early embryonic cDNA libraries (Tanaka, T.S. et al 2000, Ko et al., 2000). The set consists of 15247 unique cDNA clones which have been re-sequenced from the 3' and 5' ends identifying genes and show significant homology to known genes as well as genes unique to mouse development.

The clone set was supplied as ready to print, amplified and purified cDNA stored in forty 384 well plates. In addition the clones were supplied as bacterial cells in glycerol (glycerol stocks) formatted in 160, 96 well plates.

For the custom microarray (section 4) clones from the NIA 15K clone set were augmented by clones from the IMAGE clone set (distributed through Invitrogen, Paisley, UK) and through RT-PCR cloning procedures to form a custom clone library. For a full list of the clones used on the custom microarray and their respective source see appendix 1.

3.3.2 CREATION OF A CUSTOM CLONE LIBRARY

In order to print a custom microarray it was necessary to generate cDNA clones for each required gene in an ordered format. Genes unavailable from established libraries were cloned by RT-PCR. Figure 3.10 gives an overview of the steps used to generate cDNA clones by RT-PCR.

Firstly mRNA sequence information, for each gene to be cloned was obtained from the LocusLink database (http://www.ncbi.nlm.nih.gov/LocusLink/) and primers were designed to amplify approximately 400bp fragments of each gene (section 2.9.2). Total RNA, extracted from the mouse barrel cortex (section 3.2.3), was converted to cDNA by reverse transcription. Using this cDNA as a template, PCR amplification of each gene was carried out using specific primers (table 2.2).

The amplified gene fragments were purified and ligated to pGEM®-T vectors and transformed into DH5α E.Coli cells (section 2.14.1 – 2.14.3, figure 3.10a).

Many vectors would have failed to ligate to a PCR fragment and would have become incorporated into the bacteria without the sequence of interest, therefore PCR screens were carried out to identify cells that contained positively ligated fragments (section 2.14.5, figure 3.10b).

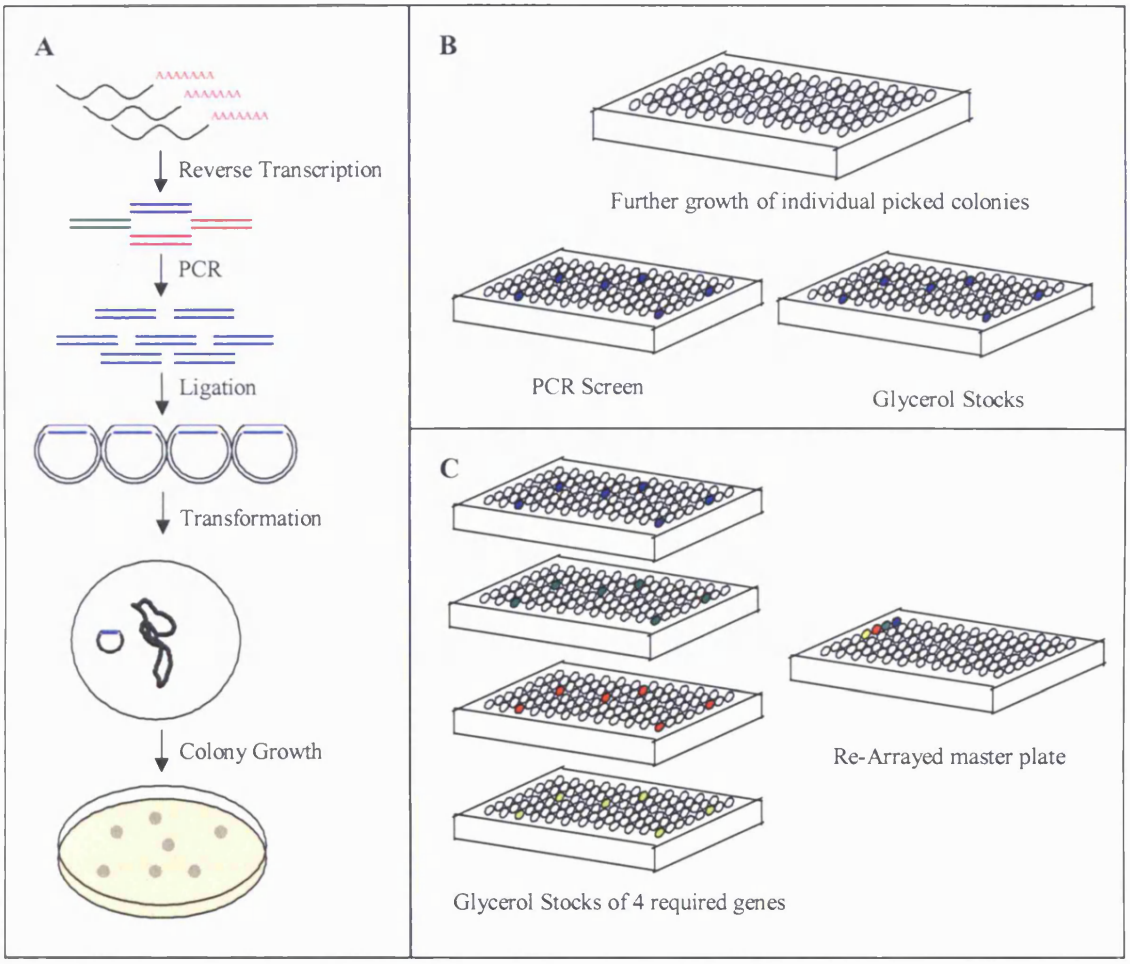


Figure 3.10 The steps involved in cloning a gene by RT-PCR prior to microarray printing.

(a) mRNA is converted to cDNA by reverse transcription. The required gene fragments are generated from the cDNA by PCR using specific primers. The gene fragments are ligated to a suitable vector allowing transformation of competent DH5α E.Coli cells. The transformed cells are spread on LB agar medium and allowed to develop until individual colonies were visible.

(b) Colonies are transferred to separate wells on a 96 well plate containing LB broth and incubated for 4 hours at 37°C. From each well, half of the culture is transferred to a new 96 well plate to make glycerol stocks while the other half is used to generate template for a PCR screen. PCR screen is used to identify wells containing positively transformed cells (BLUE).

(c) Thereby, each cloned gene required on the microarray (blue, green, red and yellow) is stored in an ordered format in bacterial cells. The PCR screen identified which well contained bacterial cells containing the gene of interest (were positively transformed) and their location on the 96 well plate. Positive glycerol stocks of each gene for the custom microarray are used to inoculate a known well on new 96 well plates containing media. New glycerol stocks are generated from these plates ensuring ease of handling and multiple processing.

To this end the DH5α transformed bacterial cells were grown overnight on a petri dish so that individual colonies formed (section 2.14.3). Single colonies were picked at random and transferred to separate wells of a 96 well culture plate containing LB broth. Typically 48 colonies were picked for each gene to ensure adequate sampling. Growth of the individual colonies was encouraged by incubating the plate at 37°C until an OD of 0.6 in the wells was reached (section 2.14.4). An equal volume of culture was then removed from each well and added to 50% glycerol in a second 96 well plate taking care to maintain the integrity of the plate layout. These glycerol stocks were stored at -80°C. The remaining cells were lysed as described and the plates spun to pellet debris (section 2.14.5). The cleared lysate was used as a template for a PCR reaction using vector specific primers. The PCR was carried out in a 96 well PCR plate and the products were visualised using a MADGE agarose gel system that maintained the 96 well layout (figure 3.11, section 2.10.2). Positively ligated cells generated PCR products of a larger size than unligated (empty vector) cells, due to the inclusion of the desired fragment, allowing their identification on the MADGE gel (figure 3.11). By using the same 96 well format throughout the process, wells containing positively transformed cells were identifiable on the glycerol stock plates.

Clones obtained from the IMAGE clone library had already been through the reverse transcription and subsequent amplification, ligation and transformation steps. It was however necessary to grow up individual colonies and PCR screen to identify positively transformed colonies of cells (as in figure 3.10).

Clones from the NIA 15K clone set had already been PCR screened and were provided as ready to use glycerol stocks again in 96 well plate format.

The last step in generating the clone library for the custom array chip was to generate master glycerol stock plates of all the required clones in an ordered format. Re-arraying all the clones together onto new 96 well plates saved space and processing time and reduced the chances of cross contamination. Each clone was given a unique identifier that linked it to a specific well on one of the plates. In total, four 96 well culture plates containing media were used to house each required clone. The wells on the plates were inoculated with their designated clones and the plates were incubated and glycerol stocks made from the growth as described. Again the glycerol stocks were kept stable by storing them at -80°C. Cloning each gene into bacterial cells meant that there was a large amount of cDNA target available for microarray printing and other work. DNA template could

be easily generated in a few hours by inoculating fresh 96 well plates from the glycerol stocks.

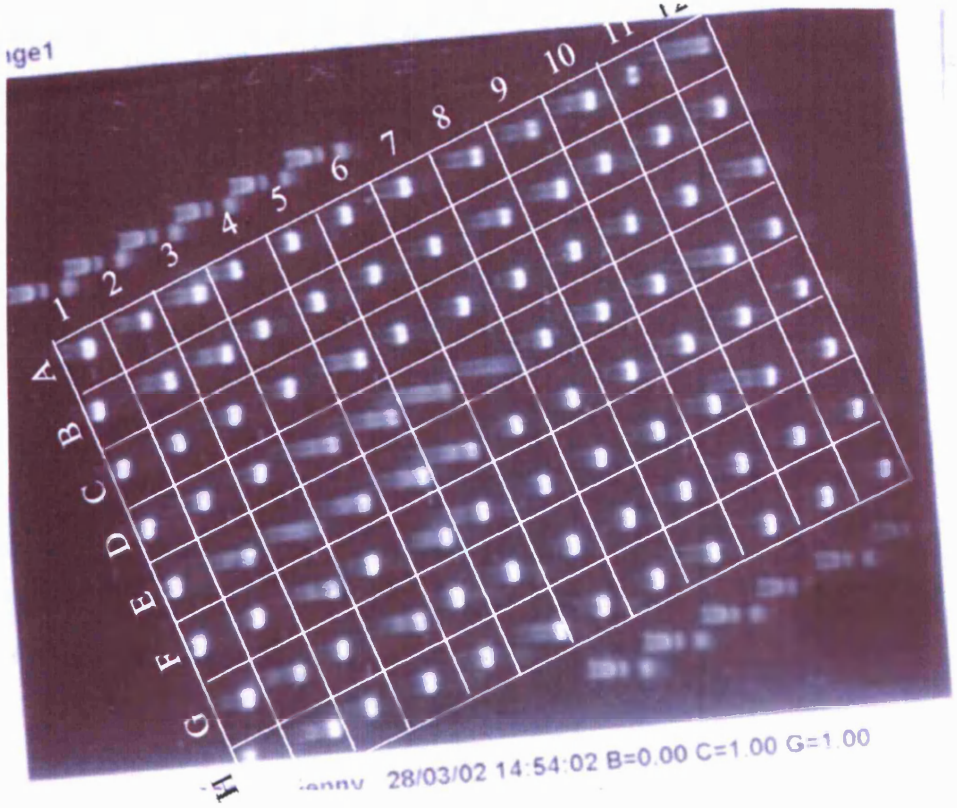


Figure 3.11 A MADGE agarose gel system used for carrying out a PCR screen during gene cloning. Following the transformation of competent cells with the gene of interest the cells are grown on suitable media on a petri dish. Individual colonies are transferred to 96 well plates containing media and grown. Half of the media is transferred to a fresh 96 well plate followed by the addition of equal amounts of glycerol and storage at -80°C. Using the remaining media the cells are lysed and PCR carried out retaining the 96 well layout. Positively transformed cells are identified on the gel by larger size fragments eg A1 – untransformed B1 transformed.

3.3.3 GENERATING CDNA FROM A CUSTOM CLONE LIBRARY

It was essential to optimise the PCR amplification of the target genes from the glycerol stocks prior to cDNA printing. Maximum cDNA product needed to be generated from the bacterial cell template to prevent over-saturation by the labelled probe during hybridisation. In other words the number of copies of the target gene spotted on the microarray must be in excess of the number of copies of the corresponding transcipt in the labelled probe (Cheung et al., 1999).

The enzyme DNA polymerase (Taq polymerase), typically isolated from the organism *Thermus aquaticus*, is used to amplify gene fragments in a PCR reaction as it is able to

withstand the high temperatures required to denature the DNA duplex (Mullis et al., 1986).

Two sources of Taq polymerase were tested to find which generated the most product during PCR amplification. In-house produced Taq polymerase, purified from over expressing cells (kindly supplied by Anna John, School of Biosciences, Cardiff University), was compared to commercially available Taq (Promega, UK) during the amplification of the GluR1 and GluR2 genes from the custom library (section 2.14). As can be seen from figure 3.12 the commercial Taq polymerase formed far more product than the non-commercial enzyme. Despite the greater cost of using commercial Taq polymerase the expense is justified as cDNA is required at as high a concentration as possible prior to microarray printing. Using this Taq polymerase, all cDNA targets for the microarray were amplified using template generated from the custom clone library (figure 3.13).

Before being used for microarray printing, the cDNA needed to be purified to remove products such as buffers and dNTPs used in PCR that may hinder hybridisation. A 96 well multiscreen filter plate, containing Sephacryl -300, provided a column purification method that was high throughput, cheap and showed minimum product retention (section 2.9.4).

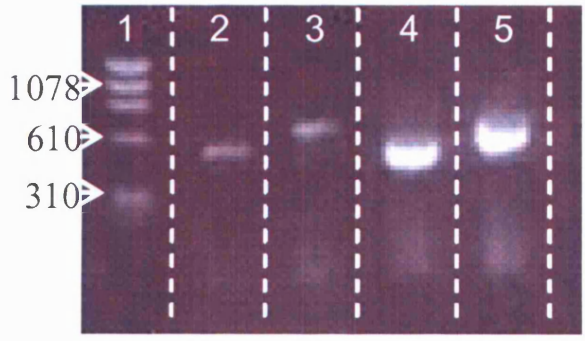


Figure 3.12 An agarose gel showing the difference in PCR product formed from identical aliquots of cloned DNA template using two sources of Taq Polymerase.

Lane 1 ØX174 size marker. Sizes shown in base pairs.

Lanes 2 & 3 PCR amplification of the GluR1 and GluR2 genes using 'in-house' generated Taq Polymerase.

Lanes 4 & 5 PCR amplification of the GluR1 and GluR2 genes using commercial Taq Polymerase (Promega, UK).

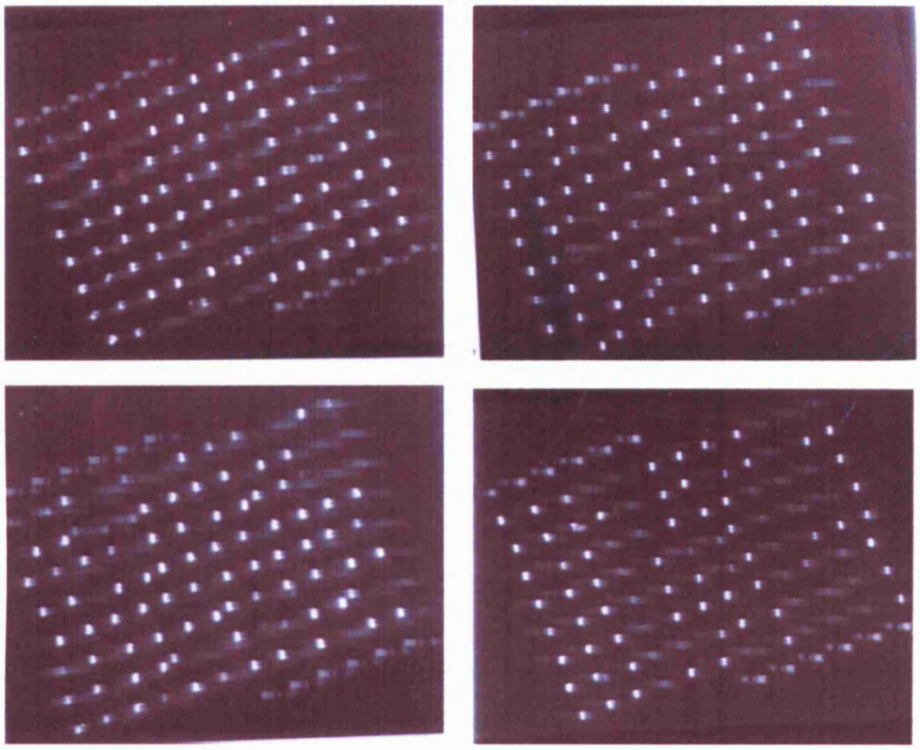


Figure 3.13 FlowMadge™ diagonal electrophoresis was used to visualise the results of the PCR amplification of each clone used in a custom microarray chip. cDNA template was generated from glycerol stocks of the bacterial clones stored in four 96 well plates.

Following purification the cDNA targets were transferred to a 384 well plate (ABgene, UK). Each clone was given a new ID corresponding to its co-ordinate on the plate. At this stage it was essential to ensure that all cloned genes contained the expected sequence and that each gene was in the correct position on the 384 well plate. Therefore all PCR fragments generated by RT-PCR and from the IMAGE clone set and 10 clones, chosen at random on the 384 well plate, from the NIA 15K clone set were sequenced (see figure 3.14, section 2.9.5). 42 out of the 52 clones sequenced were in the correct position and contained the expected inserted sequence with the 10 failing to provide a readable trace. Finally an equal volume of DMSO was added to each well and the plate was stored at -20°C.

Random samples were taken from the 384 well plate, pooled and the concentration of the cDNA was determined spectrophotometrically (section 2.11) to be 205µg/ml. This concentration is within the optimum range for spotting on microarrays (Yue H. et al., 2001)

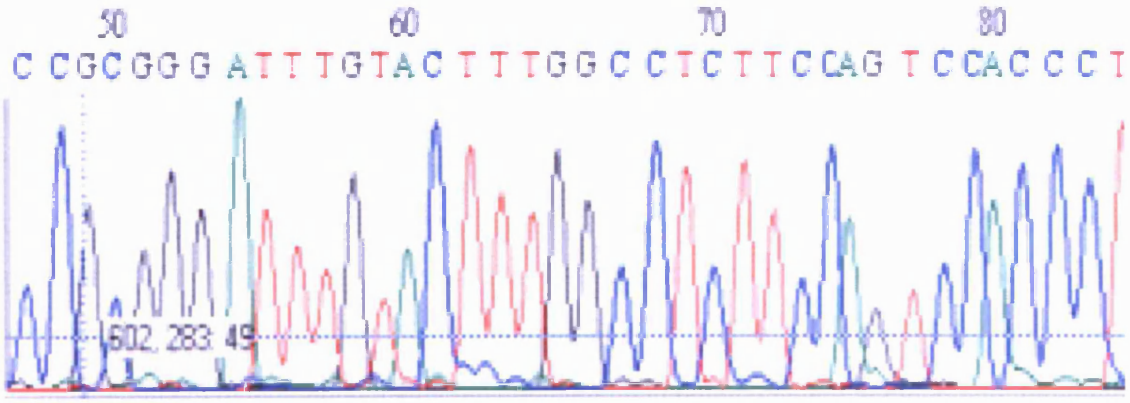


Figure 3.14 Sequence trace typical of reactions carried out on selected cDNA targets.

Sequencing of cDNA products prior to microarray printing ensured correct positioning within the 384 well plate and also that gene fragment cloning had been successfully carried out.

An example of the sequence trace as seen on the BioEdit (Version 4.8.10) sequence alignment editor software. The sequence is generated from bacterial plasmid DNA transformed with a PCR amplified mouse cDNA fragment of the GluR1 gene. Using blast 2 sequences program at http://www.ncbi.nlm.nih.gov/BLAST/. The sequence generated from the bacterial clone shows 100% homology to the expected sequence, in this case the mRNA of the mouse GluR1 gene.

3.3.4 MICROARRAY PRINTING

Two types of microarray were used in this investigation. The small-scale microarray slides (section 4) were printed using a Flexys™ High Density Arrayer (Genomic Solutions, Cambridge, UK) while the high density slides (section 5) were created on a SpotArray™72 (Perkin Elmer™, UK).

Both printers deposited cDNA from 384 well microtitre plates onto the microarray slides (see section 2.15.2.4 and 2.15.3.2). The operator was required to input the number of 384 well plates, plate contents and desired printing parameters (eg number of replicates) using the print control software. An input file (.csv) was created in which each well on the 384 well plates was given a unique identifier, in addition to further information about the cDNA clone. Following printing an output file was generated that gave each spot a unique identifier relating to its position on the array. This position was determined by the meta grid and sub array location of the spot (see figure 3.15). The controlled printing ensured that each spot could be identified using this unique 4 co-ordinate address. The output file merged the data from the input file, using the 384 well address field common to both files, so that each spot could be traced to its respective cDNA clone.

The printers used a head which held 48 pins in place. The Flexys™ High Density Arrayer used solid titanium pins to spot attached cDNA. In contrast the SpotArray™72 used TeleChem Stealth Quill pins which used capillary action to draw cDNA into the tip. The quill system was able to print up to 100 times from a single uptake, cDNA being deposited on the microarray slide by a tapping action. In contrast the titanium pins would print once and would then return to the 384 well plate to pick up fresh cDNA. The advantage of the quill system was its ability to form very small spots on the slide resulting in slides with far more spots (approximately 15000 spots at 10 pixel diameter compared with 2352 at 23 pixels diameter).

To prevent carry-over, and in the case of the quill pins, prevent clogging, before, during and after printing the pins were cleaned with a wash system incorporated into the printers. The quill pins were cleaned using high pressure water jets and dried using a vacuum system. The solid pins were cleaned by using a brushing bath and dried using a halogen lamp (see section 2.15.2.4 and 2.15.3.2).

Glass microscope slides provided the solid substrate for the microarray offering transparency, rigidity, resistance to temperature extremes and also allowing chemical modifications to the surface. Chemical modifications are carried out on the slides to increase target binding and decrease non – specific background binding (Holloway et al., 2002). Poly-L-lysine coated slides are the choice for arraying on a budget but the general consensus in the arraying community is that CMT-GAPS™ coated slides (Corning, NY, USA) although more expensive give much improved results (Hegde et al., 2000). These slides are coated with Gamma Amino Propyl Silane which offered high DNA binding capacity. This high affinity binding occurs via ionic interactions between free amine groups of the GAPS and the negatively charged phosphate backbone of the cDNA target. CMT-GAPS coated slides were used throughout this project. Both printers were housed in a suite where the temperature was kept at an optimum 22°C. The atmosphere was filtered to remove dust particles and constantly monitored for humidity. Printing was only carried out when the relative humidity was 45-50% (Hegde et al., 2000).

Following printing the slides were dried at 80°C for 3 hours before cross-linking. Slides were stored under vacuum in a light safe box.

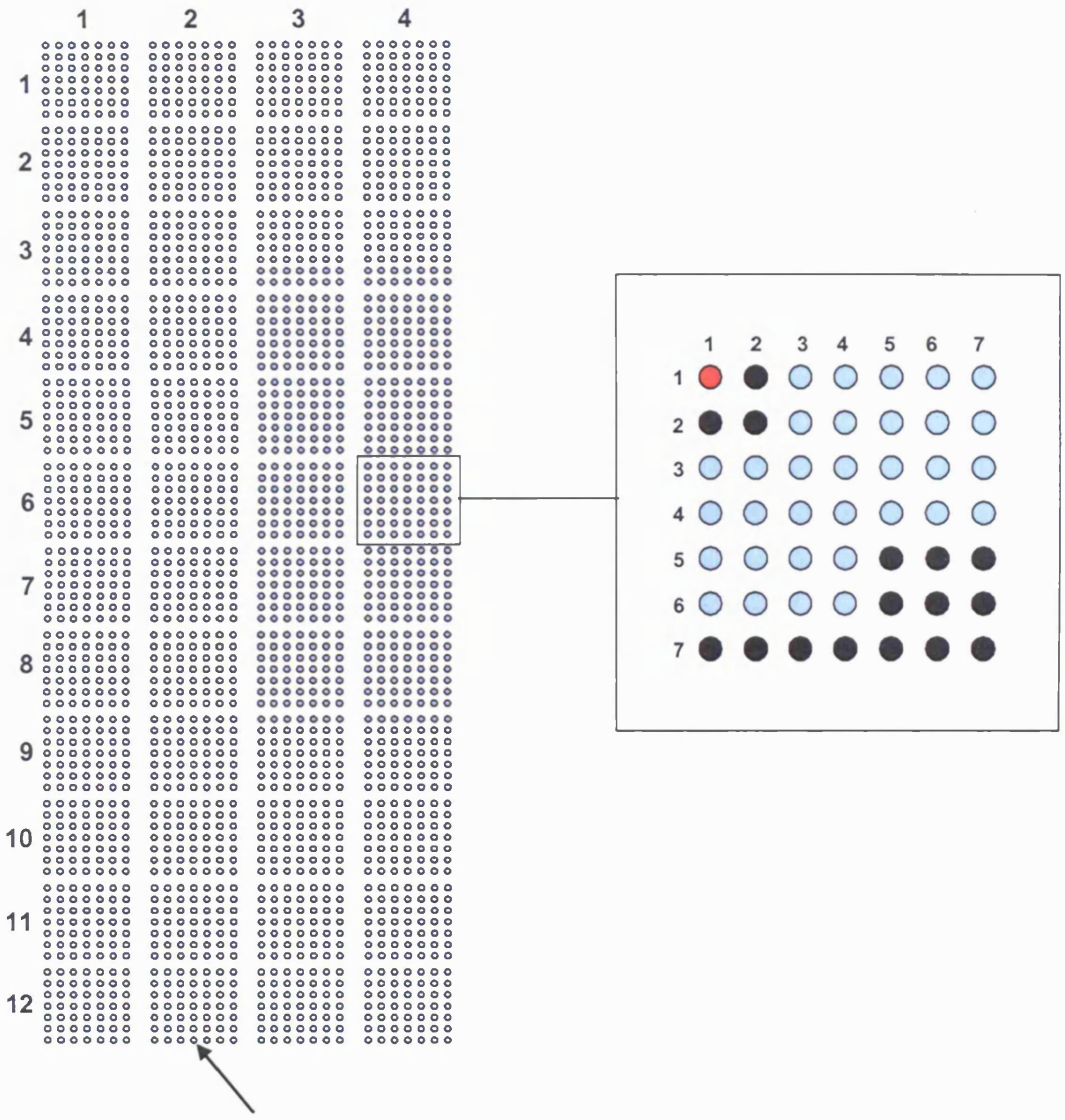


Figure 3.15 The layout of spots on a custom microarray chip.

Each cDNA clone on a 384 well plate was printed 4 times by a Flexys™ High Density Arrayer (Genomic Solutions, Cambridge, UK). The array is made up of a meta grid consisting of 12 x 4 sub-arrays of 49 spots (7 x 7). Each spot had a unique identifier given by Meta Row, Meta Column, Sub-array Row and Sub-array Column eg 1,12,4,7 (arrow). A csv file generated by the arrayer software ties the spot identifier to its position on the 384 well plate. Each sub-array consisted of landmark spots (red), blank spots (black) and cDNA clones (blue)

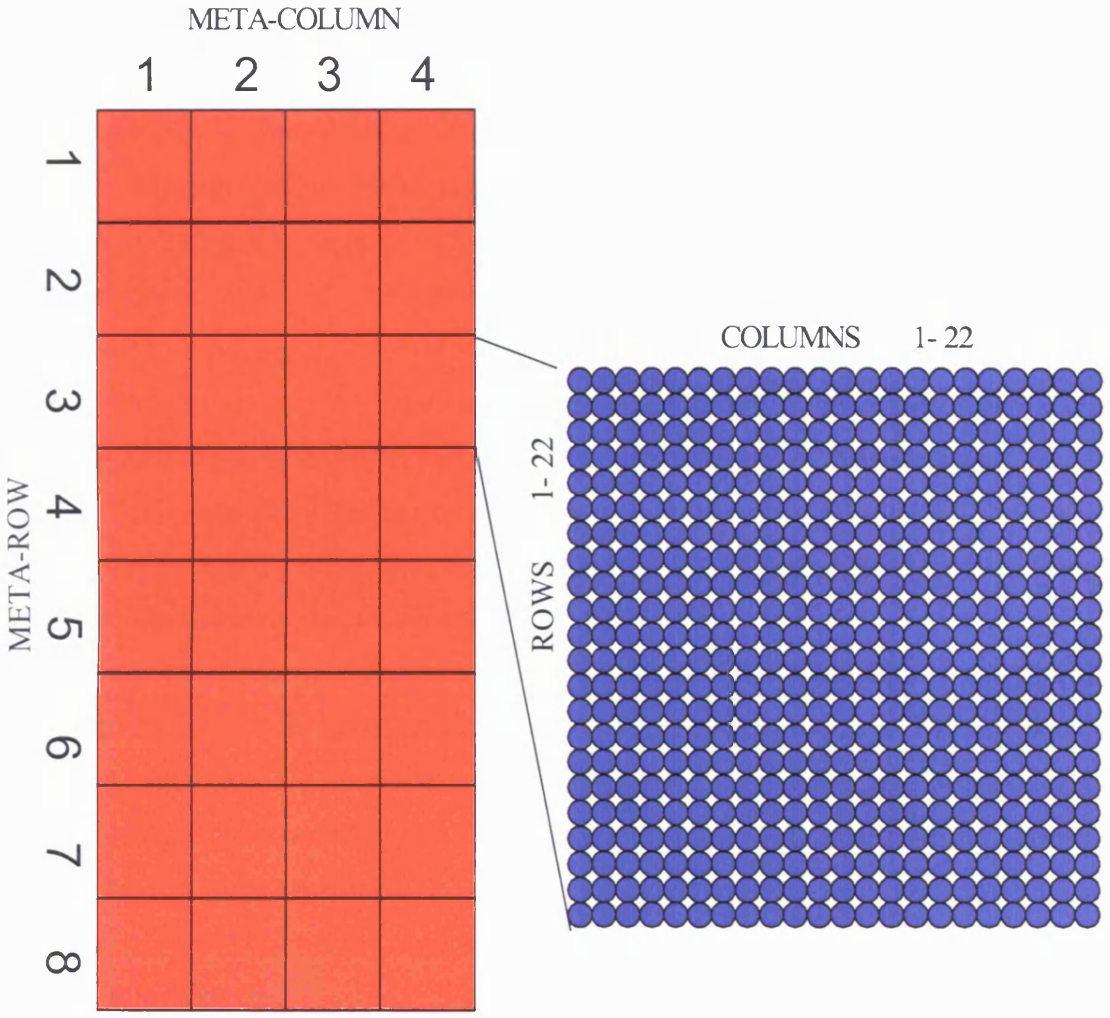


Figure 3.16 The arrangement of spots on a NIA 15K clone set microarray.
The spots are printed in sub-arrays (blue) made of 22 rows and 22 columns. Each sub-array is part of a meta-grid (red) which consists of 32 sub-arrays arranged on the microarray in 8 rows and 4 columns. Thus any spot on the array can be identified using 4 pieces of information (Meta-Row, Meta-Column, Row and Column (of the sub-grid). In appendix 3 this information is tied to a unique identifier that provides the identity of a specific clone from the NIA clone set. Thus spots on the microarray can be related to a gene.

A technique was required that allowed the quality of printed microarray slides to be assessed prior to hybridisation with probe. SYBR® Green staining (section 2.15.5) was used to visualise the DNA deposited at each spot on the microarray and to evaluate the overall efficiency of the printing process (Battaglia et al., 2000).

Following each print run, slides were randomly picked from batches for SYBR® Green staining (typically one slide for every 10 printed). SYBR® Green is a fluorescent dye which specifically binds to DNA and allows the visualisation of deposited target cDNA spots on microarray slides. This proved to be a cheap and convenient way of assessing the array printing process prior to hybridisation with probe, the only disadvantage being that the chosen slides had to be discarded following the stain. The stained slides could be visualised by scanning at 600nm using a GeneTAC™LS-IV scanner (Genomic Solutions, UK). Stained slides were qualitatively checked for spot morphology, whether grids conformed to the expected print patterns and the overall equality of deposition.

Figure 3.17 shows the difference between a high density (a) and low density (d) microarray. Clearly visible and showing good morphology are DNA spots in the grid formation specified by the print control software.

For the custom arrays water blanks were incorporated at known positions in the microtitre plates to control for the possibility of carry-over contamination, as a result of poor pin cleaning. Figure 3.17d shows the position (enclosed by white area) corresponding to water blanks. As expected no cDNA spots are visible in this area therefore carry-over contamination has not occurred.

Figures 3.17b and c show the results of two stains carried out on high density microarray slides from separate print runs. Figure 3.17a shows a good quality slide with clearly visible grids of DNA spots. Figure 3.17c shows an example of a poorly printed batch. In this case the pins failed to come into contact with the DNA resulting in large empty patches on the slide with no spots. The SYBR® Green stain technique enabled low quality batches to be rejected saving precious probe from being wasted and is a valuable method for detecting gross defects in the microarray printing process.

More subtle problems with printing came to light following hybridisation. During this stage of the process it was procedure to inspect all spots on an individual basis (to check spot morphology and spot finding (see section 3.6.2). Examples of typical spot morphologies are given in figure 3.18. Well printed spots show high circularity, equal fluorescence and good spacing between adjacent spots (figure 3.18a). Figure 3.18b and c show abnormal spot morphologies caused by less than optimal printing conditions. For best results slides should be printed at a temperature of 22°C and 42% relative humidity (Hegde et al., 2000). Departures from these conditions can have serious consequences on the print quality, therefore printing was carried out in a temperature controlled room, with

constant monitoring of the humidity. The effect of printing at a higher humidity is shown in figure 3.18b where spots have become diffuse and merged. This effect was more apparent in the high density arrays where there is little space between spots.

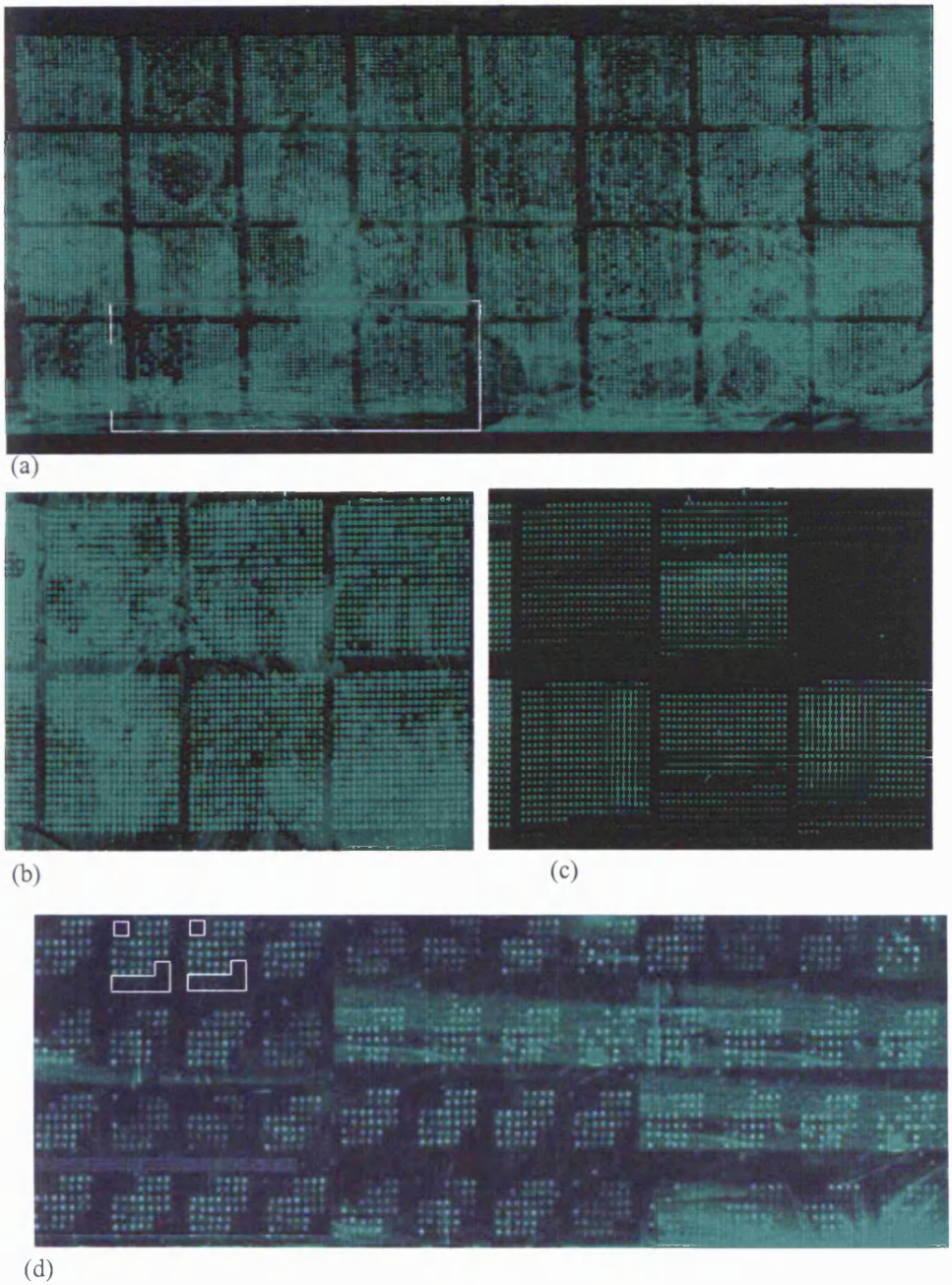


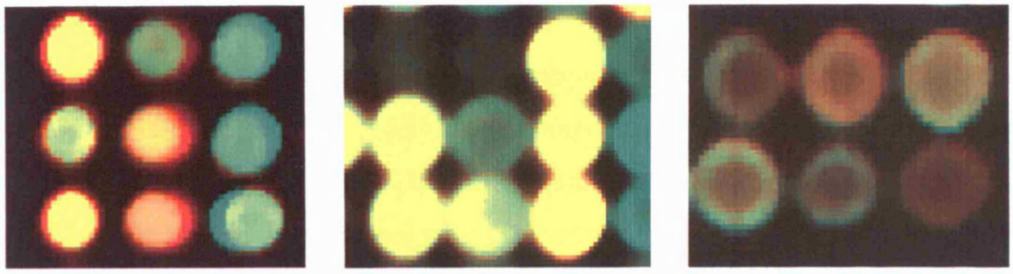
Figure 3.17 Examples of SYBR® Green staining of microarray slides to evaluate print quality

(a) A high density array printed using SpotArray™72 (Perkin Elmer™, UK).

(b) Magnified view of 3 blocks from (a) showing expected grid pattern.

(c) Poorly printed slide (note dark patches with missing spots).

(d) Low density array printed using Flexys™ High Density Arrayer (Genomic Solutions, Cambridge, UK). Stain confirms expected print pattern and no carry-over contamination (white area).



(a) (b) (c)

Figure 3.18 Common morphological features of microarray spots.

(a) Well printed spots with good circularity and uniformity.

(b) Effects of humidity on spot morphology. Adjacent cDNA spots have merged and the data collected is compromised.

(c) Doughnut shaped spots.

The TeleChem Stealth Quill pins worked by drawing a volume of cDNA into the tip (approximately 250nl) and tapping each required slide to deposit the cDNA. About 100 slides can be printed from a single uptake. It was therefore important to add 50% DMSO to the cDNA target prior to printing to prevent sample evaporation from the tip during lengthy print runs (McQuain et al., 2003). The hygroscopic nature of DMSO combined with drying at 80°C for 3 hours (as opposed to a longer drying time) also reduced the formation of 'doughnut' spots (figure 3.16 c). These features are thought to occur when rapid evaporation from the edges of spots draws liquid from the centre, by capillary action resulting in localised depletion (Deegan et al., 1997). Other theories suggest that crystallisation promotes formation of a complex in the spot centre preventing probe hybridisation. The DMSO had the added effect of denaturing the double stranded cDNA therefore encouraging binding to the CMT-GAPS coated slides and making the cDNA available for hybridisation. The occurrence of diffuse spots was rare but the doughnut shaped spots appeared more frequently. Doughnut shaped spots were included in further analyses as contamination had not occurred and both Cy3 and Cy5 fluorescence would have been equally affected by these features. Diffuse spots were excluded from any analysis.

3.4 HYBRIDISATION

Hybridisation can be split into three parts; pre-hybridisation (blocking), hybridisation and washing.

For all hybridisations probes were combined and concentrated using an Eppendorff Concentrator. The concentrated probes were re-dissolved in a suitable amount of hybridisation buffer (typically 30µl for a manual hybridisation and 110µl for the automated system, see section 2.16.1). The composition of the buffer was important to the success of the hybridisation. The buffer contained 50% formamide which had the effect of lowering the melting temperature (Tm) of the probe. This meant that 50% of the probe was denatured at a temperature of 42°C rather than approximately 70°C in aqueous buffers (Hegde et al., 2000). Consequently, hybridisations could be carried out at 42°C significantly lowering the chances of probe evaporation.

Following the addition of hybridisation buffer it was necessary to block certain sequences common to many genes to prevent non-specific cross hybridisation. To this end Poly dA^{30} and mouse COT1-DNA were added to the probe solution to block Poly-dT sequences formed during cDNA synthesis and repetitive genomic elements such as B1, B2 and L1 family members (Hegde et al., 2000, Landegent et al., 1987 and Lengauer et al., 1990). The solution was then heated to denature and snap cooled on ice before application onto the microarray.

The process of microarray pre-hybridisation (blocking) aims to reduce the amount of non-specific bonds formed between the probe and the glass slide during hybridisation. A pre-hybridisation of the microarray in a 1% solution of Bovine Serum Albumin (BSA) was sufficient to block the interactions of probe and free amine groups on the slide (Hegde et al., 2000).

Here the manual method of hybridisation was used in which a coverslip is lowered over a microarray slide to which labelled probe has been applied. For the manual hybridisation the coverslip had to be lowered very carefully to ensure bubbles did not form. If bubbles did form they were virtually impossible to remove and no hybridisation could take place in that area of the slide resulting in a loss of data. This process was practiced, using buffer only, prior to hybridisation with probe. To prevent the buffer evaporating at the edges of the coverslip the hybridisation was carried out in a pre-heated (42°C), humidified chamber within a light proof oven (section 2.16.3.1). Care was taken to prevent the slide from touching the water in the chamber as a loss of probe would have occurred. In an effort to prevent contamination of the slide, by dust and other contaminants, powder free gloves were used at all times and the slide was treated with an anti-static gun during handling. As the probe is light sensitive, exposure to light was minimised by carrying out all stages of the procedure as quickly as possible.

For automated hybridisation, although the slides are handled less, it was still important to reduce the chances of dust contamination. The application of probe to the array was simplified by using an inlet port on the hybridisation chamber eliminating bubble formation. The procedure was quicker and simpler than the manual hybridisation (section 2.16.3.2).

Following hybridisation the microarray slides were washed to remove hybridisation buffer components and loosely bound probe. To wash, the microarrays were incubated, at 42°C in pre-warmed high salt content buffers followed by low stringency buffers and filtered water (section 2.16.3.1). For manual hybridisations the washes were carried out in light-safe Coplin jars while the automated system washed the slides *in situ* from stock buffer bottles on the machine (section 2.16.3.2)

3.5 SCANNING

Microarray experiments use a high-resolution fluorescent scanner to visualise the fluorescence generated by the fluorophore labelled probes. Early microarray work used scanners that were manufactured in-house (Schena et al., 1995, Cheung et al., 1999). The boom in the use of array technology has seen a host of companies now mass producing scanners dedicated to the analysis of microarrays eg Perkin Elmer®, MA., USA, Agilent Technologies, USA, Axon Instruments, CA, USA, Bio-Rad Laboratories, CA, USA and Genomics Solutions®, Cambridge, UK. This project utilised the Genomic Solutions® LSIV scanner owned by the Cardiff University School of Biosciences.

The scanner houses a laser that delivers a beam of light at the microarray slide. Coverage of the entire microarray is achieved by moving the excitation beam in the x- and y- axis over the microarray slide. Excitation filters exclude light of certain wavelengths enabling the system to be tuned to the excitation wavelengths of the fluorophores being used (Lyng et al.,2004).

Following excitation the moving head collects the light emitted by the fluorophores. Seven collection filters in the head exclude reflected and scattered light from the excitation beam, ensuring only light of particular wavelengths, namely the emission wavelengths of the fluorophores, are collected. On collection a photo-multiplier tube (PMT) is used to detect and quantify the emitted light. Emitted light is of a very low intensity therefore the PMT amplifies the signal by converting the photons of emitted light to electrons. The PMT multiplies the number of electrons by acceleration in a

voltage gradient generating a measurable electrical signal in proportion to the number of photons. The Genomic Solutions® LSIV uses 16-bit digitization to convert the continuous analog signal from the PMT to a digital signal with 65536 discrete intensity values. This provides the scanner with the necessary amplitude resolution to detect and resolve low and high intensity pixels (or high and low expressed genes) simultaneously (Schermer, 1999).

One important fact about the Genomic Solutions® LSIV scanner is its high spatial resolution. A scanned array image is made up of individual pixels. The size of each pixel is dependent on the diameter of the light beam when it reaches the microarray. The narrower the beam of light, the smaller the pixel size and the better the spatial resolution of the image (Schermer, 1999 and Chen, Dougherty and Bittner, 1997). Most scanners developed for microarrays, including the Genomic Solutions® LSIV, have a resolution of 10μm ie each pixel is 10 x 10um. This resolution is enough to visualise approximately 1750 separate images within an array spot 150μm in diameter. Pixel size is fully adjustable up to 150μm for rapid scanning of agarose gels and microarray previews (section 2.17).

3.6 IMAGE ANALYSIS AND DATA EXTRACTION

Following hybridisation microarray slides are scanned and the image generated is analysed. This section aims to optimise the steps involved in the analysis of a microarray image and the extraction of data using a Self-Self hybridisation.

3.6.1 SELF-SELF HYBRIDISATION

The aim of a typical microarray experiment is to identify differentially expressed genes from two distinct RNA populations. During the optimisation of the microarray process a wealth of information can be gained by carrying out a Self- Self hybridisation (Yang IV et al., 2002). Here, aliquots from the same RNA sample are labelled with Cy3 and Cy5 dyes and hybridised to the same microarray.

The data generated from such an experiment provides information about the variation inherent in the array process. More importantly, the way the data is extracted from the image can be optimised as the results are not complicated by differential gene expression. In the following section the processes of image analysis and data extraction are optimised using Self – Self hybridisations on a custom microarray chip and a 15000 spot microarray slide.

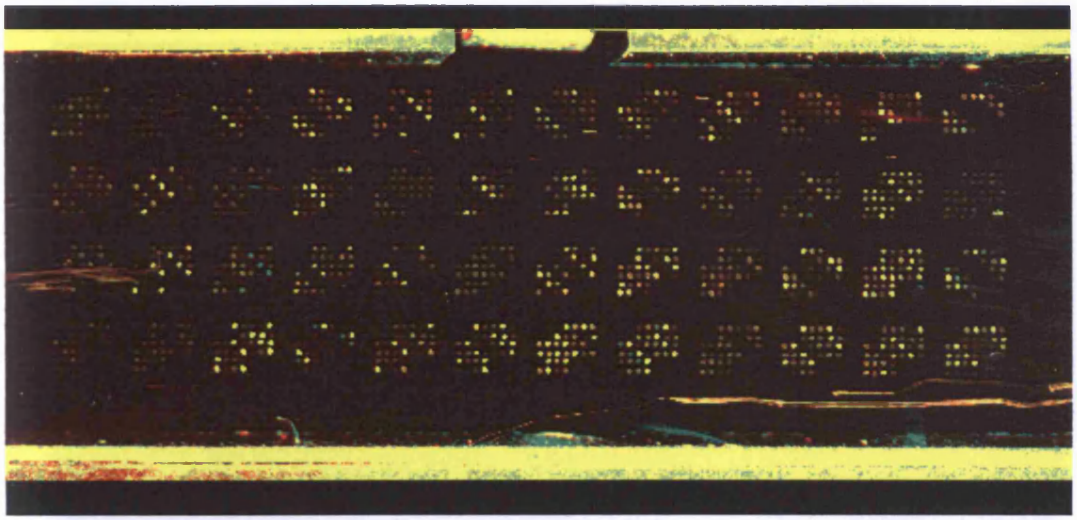
For the custom microarray 10μg aliqouts of pooled, total RNA (section 3.2.3) were labelled with Cy3 and Cy5 and hybridised to the custom slide (section 2.16.3.1). The 15,000 spot microarray was probed with amplified RNA (aRNA) generated from pooled total RNA (section 3.2.3). Two 1μg aliquots of aRNA from the same amplification reaction were fluorescently labelled and hybridised to a 15000 spot array.

3.6.2 SPOT FINDING

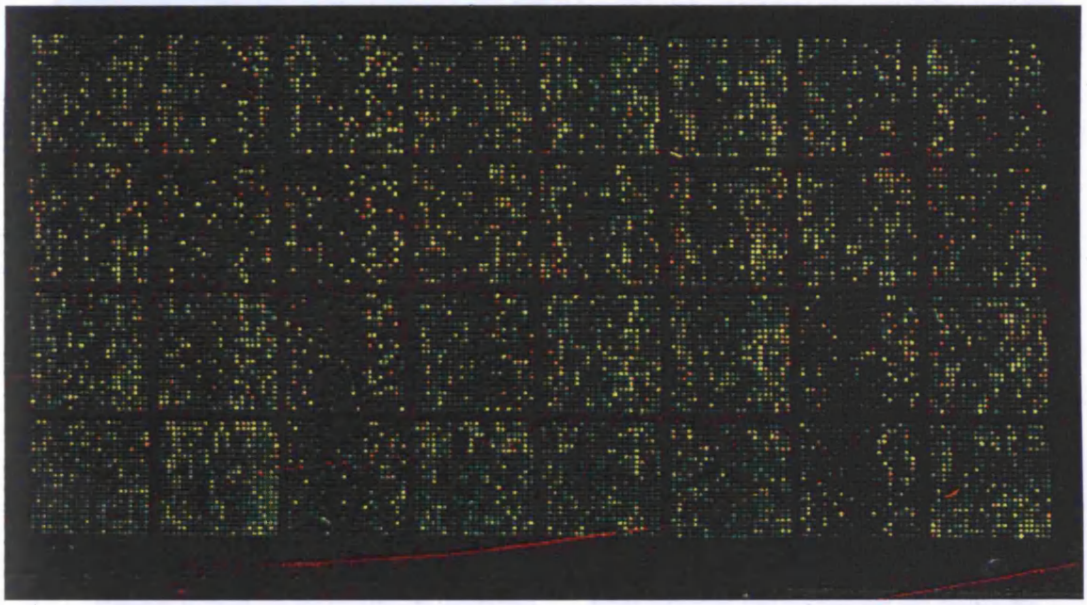
Following hybridisation the microarray slides from the Self-Self hybridisations were scanned, using a GeneTAC™LS-IV at the appropriate wavelengths as described (section 2.17). The overall yellow appearance of the arrays (see figure 3.19) is indicative of equal amounts of Cy3 and Cy5 labelled probes. During scanning each fluorophore is assigned a false colour to aid analysis. The convention is to use green for Cy3 and red for Cy5 therefore equal amounts of probe appear yellow (Yang YH(a) et al., 2001). Scanning captured an image of the slides that included spots and background. The image generated was broken down into pixels, the number of pixels dependent on the resolution of the scanner (section 3.5). Images were imported into Imagene™ (Biodiscovery, Inc. CA, USA) microarray analysis software from the scanner as TIFF files. The relatively large size of these files (approximately 17mB per channel) and the subsequent processing involved, meant that a PC with a recommended 512mB RAM operating on a Microsoft® Windows® XP platform was required to support the software. To accurately quantify the amount of fluorescence at each spot the software needed to know the precise location and boundary of each spot (Yang YH et al., 2001). This process is known as spot finding.

Prior to spot finding the user had to import the output file generated following printing of the microarray (section 3.3.4). Generated by the print control software this file gave each spot a unique identifier which allowed it to be traced back to its respective position on the 384 well plate containing the cDNA used to print the array. Importantly, the file also provided the meta grid and sub array layout of the spots (figure 3.15) in addition to the expected spot sizes. Using this file the software knew the expected number and layout of spots on the array.

The user then had to mark on the image the spots at the extremity of the meta grid firstly, followed by those in the sub grids. The software would then place grids over the anticipated position of the spots (figure 3.20). Within each square of the sub grids was a circle that should have contained the array spot.



(a)



(b)

Figure 3.19 The images generated following the scanning of two microarray slides following Self – Self hybridisations. The overall, yellow appearance of the images is the result of equal amounts of fluorecence from the Cy3 (green) and Cy5 (red) labelled probes. The false colours are generated by the scanner software to avoid confusion when analysing the image.

(a) Cy3 and Cy5 labelling reactions were carried out on 10µg aliqouts of pooled Total RNA. The probes were hybridised to a custom microarray chip and the slide scanned.

(b) Cy3 and Cy5 labelling reactions were carried out on 1µg aliqouts of amplified RNA from the same amplification reaction. The probes were hybridised to a microarray chip, printed using the NIA 15k mouse clone set and the slide scanned.

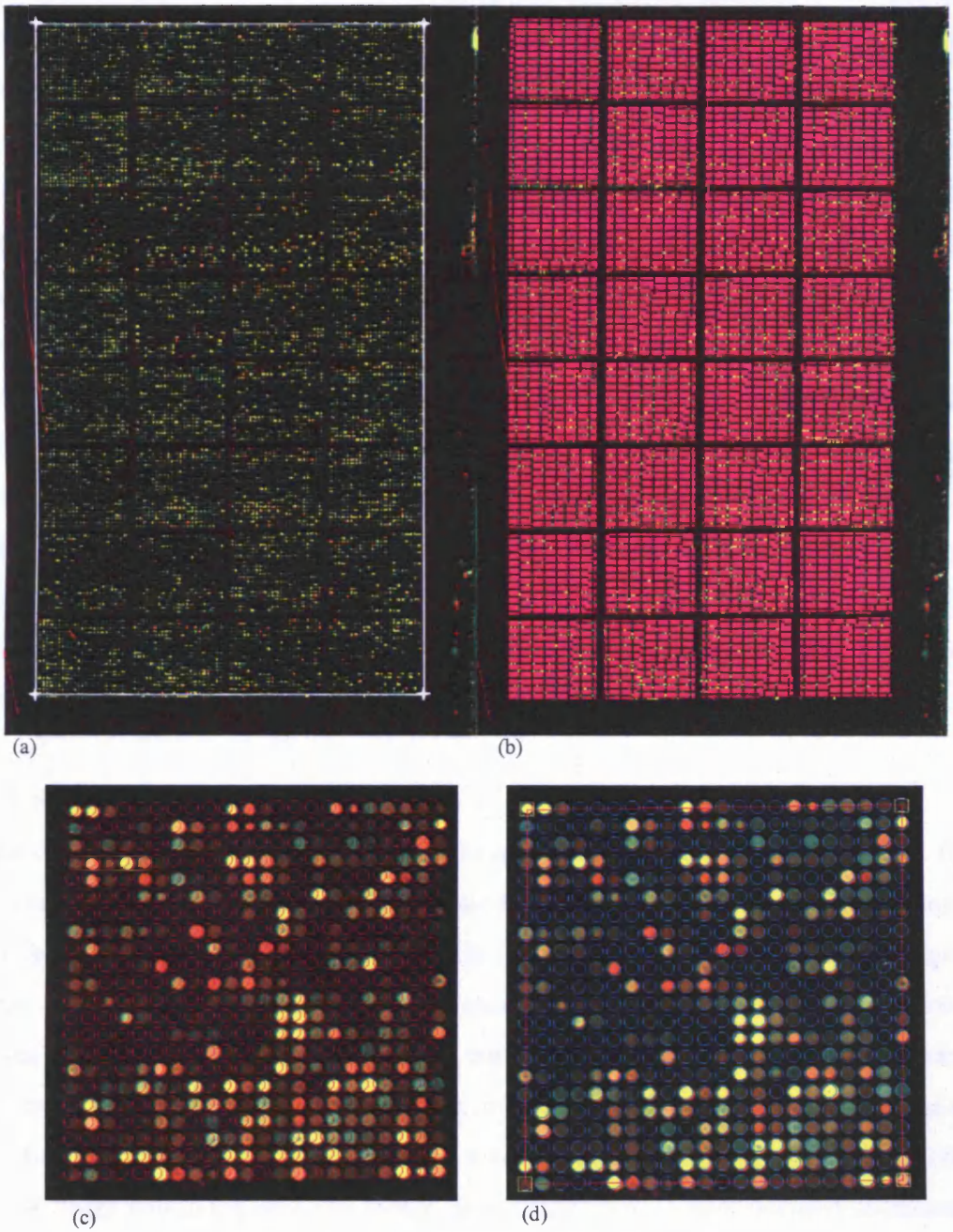


Figure 3.20 a- d the process of spot finding (shown for a 15000 spot array).

(a) Original image of a scanned microarray slide following hybridisation. As seen using Imagene™ Software. Process of gridding begins by identifying corner spots on meta grid (white stars and lines).

(b) MetaGrid is overlaid onto array image. MetaGrid is broken down into 12 x 4 sub-arrays.

(c) Imagene™ Software places red circles on expected spot positions as defined by Quant Array text file. Minor discrepancies in printing lead to most circles being misaligned with actual spots.

(d) Grid around each sub-array is manually adjusted so that circles align perfectly with spots.

The initial grid placement carried out by the software always failed to perfectly place circles over the printed spots (see the red circles in figure 3.20c). This is due to small

inconsistencies that occurred during printing where a difference of a few microns in spot printing had a knock-on effect in grid overlay. To counter an automated function for spot finding following initial grid overlay was used in Imagene™ Software. The algorithm was intended to completely automate the task of grid adjustment and perfectly place circles over the spots. However, the function failed and many circles were still misaligned. The algorithm could be adjusted (using a sliding scale in pixels) so that the software searched for spots at greater distances from their expected positions. Though the software was afforded this greater flexibility many circles failed to overlap spots. The sub grids had to be manually adjusted so that approximately 80% of the area of all spots were within the circle generated by the Imagene™ Software. Only then would the automated function adjust the spot area to completely cover the spots (see figure 3.20d). This method however was still superior to early image analysis software packages (Eisen et al., 1998, http://rana.lbl.gov/EisenSoftware.htm) in which the placement of circles over every spot had to be carried out individually. A robust, automated spot finding program would be a great improvement over the semi-automated process used and is the focus of research for a number of groups (Jain et al., 2002, Ahmed et al 2004).

3.6.3 SEGMENTATION

Once the Imagene™ Software had placed circles around the spots and the spot finding had been checked the process of segmentation was carried out. The role of segmentation is to accurately separate background pixels and contaminants on the array image from signal pixels (Soille, 1999). This was necessary because Imagene™ Software drew circles of an expected size around each of the microarray spots. If the software had used this circle area as the spot boundary, then, in the case of small spots, spot pixels would have been contaminated with background pixels. Also any contaminating artefacts such as dust flecks within the spot area would have contributed to spot intensity measurements and given a distorted measure of spot intensity (Yang YH et al., 2001).

The segmentation process combined spatial and intensity information about each spot and its immediate background and a trimmed measurement method was then used to remove contaminating pixels. Firstly, the distribution profiles of the Cy3 and Cy5 intensities of all pixels, within the signal and background areas for every spot on the array are plotted. A percentage of high and low intensity pixels were then removed from the analysis. By removing high intensity pixels from the signal areas, dust and other artefacts could be removed. By ignoring lower intensity pixels the software now rejected background pixels

contaminating the signal area on smaller spots. This exclusion applied to every spot on the array and by selecting any spot the area excluded could be visualised. Manual segmentation required the user to input the percentage of high and low pixels to exclude. Automatic segmentation was carried out using a 'robust patent pending statistical approach' built into the Imagene™ Software that calculated the segmentation parameters for the user. Both methods were tested using the Self-Self hybridisation carried out on the custom microarray slide. For a variation on the histogram segmentation method that statistically assesses each pixel to assign them as signal or background see Chen et al., 1997.

The automatic segmentation tended to use less strict parameters resulting in many spots showing signs of contamination (figure 3.21b) whereas the manual approach appeared to remove all contaminating pixels (figure 3.21c). This contamination is evident in figure 3.26 which are plots of the mean pixel intensities against the median pixel intensities, for every spot on the Self – Self array, using both manual and automated segmentation. As a rule the median is less sensitive to outliers than the mean. If spots were free from outlier pixels (contamination) then the pixel mean and median for the spots should be equal (Tran et al 2002,). Manual segmentation (figure 3.22a) results in less spread of the data than automated segmentation (figure 3.22b) and all points tend to lie on the line y=x suggesting that the mean and median are more equal in the data generated using manual segmentation. This suggests that manual segmentation is a better method of removing contaminating pixels than automated segmentation. Therefore manual segmentation is used in all further analysis.

One thing that must be considered when using manual segmentation is that too high a proportion of pixels are not rejected. If this occurs in spots with no contamination then sampling is reduced and the accuracy of the data is diminished. However this can be controlled by setting a threshold (usually approximately 20% of pixels) for removal. In such a way contaminants can be removed without affecting the sampling rate. Following segmentation any spots that remained contaminated were removed from analysis by manually flagging them as poor quality.

3.6.4 DATA EXTRACTION SOFTWARE

Following gridding and segmentation Imagene™ Software quantifies the amount of fluorescence in each spot and immediate background and presents the data in a .txt file. As two flurophores (Cy3 and Cy5) were used to label the RNA samples under

investigation, two .txt files were produced for each array, one for Cy3 and one for Cy5. Each file contained information relating to the positioning of each spot on the array spot quality information (such as spot diameter, whether the spot is flagged as good or bad) and the quantified fluorescence data.

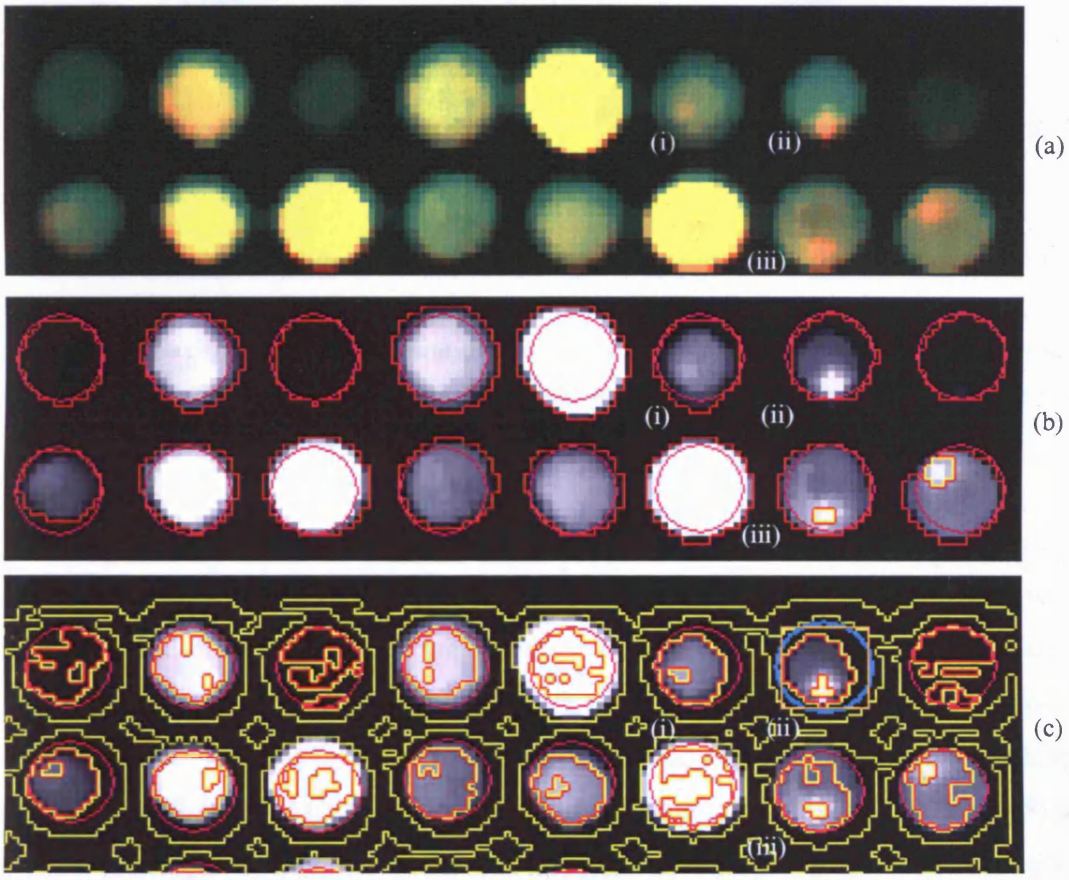


Figure 3.21 a, b and c. Segmentation process as carried out using Imagene™ Software.

(a) Composite image of a selection of spots viewed before gridding . Note smaller area of spots (i and ii) and dust specks (red marks) on spots (i, ii and iii).

(b) Using the automatic segmentation facility in Imagene™ Software the image is broken into signal pixels (inside red jagged circles) and background pixels (everything outside the red circles). Regions of contamination within the spots (shown by yellow border) are excluded from quantitation of signal intensity. Note that segmentation has failed to remove contamination from spots i and ii and background pixels are included in the spot area.

(c) Segmentation is carried out manually. The segmentation parameters used are more stringent than those for automatic segmentation and contaminants (dust and background pixels) are successfully excluded from the spot area.

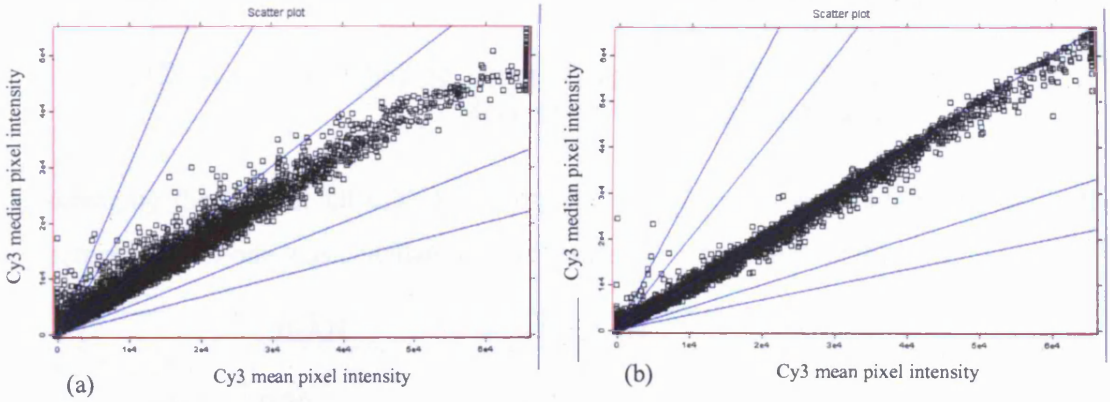


Figure 3.22 a and b. A comparison of the data output quality using automated (a) and manual segmentation (b) to separate signal pixels from background / contamination pixels. Both graphs plot the mean (y-axis) against median (x-axis) of all Cy3 spot intensities from a Self – Self hybridisation on a custom microarray (see section). The greater spread of the data in (a) and the fact that in (b) all spots cluster on the line y=x is reflective of cleaner data generated using manual segmentation.

In order to analyse microarray data GeneSight® software suite (BioDiscovery, Inc.,CA), Imagene®'s image analysis program, was used. At the front end of the software was a simple drag and drop menu which was used to combine the two .txt files from a single microarray image. The software paired the Cy3 and Cy5 data by spot location to produce a single .txt file. This file provided measures of the amount of fluorescence in each spot for both channels (Cy3 or Cy5), background fluorescence for each spot (Cy3 or Cy5) and information on spot quality. In addition to the quantified intensity for each fluorophore the software calculated the ratio of Cy3 intensity to Cy5 intensity.

3.6.5 METHOD OF QUANTIFICATION OF FLUORESCENCE

The output file produced by GeneSight® provided three measures of the spot fluorescence; the mean, median and mode of the total pixel intensity. In an effort to reduce variation it was important to find which measure provided the more robust quantification of fluorescence.

To this end the data generated from the Self-Self hybridisation on the custom microarray slide was used. The Cy3 / Cy5 ratio for each spot using either the mean, median or mode was provided in the results file. As each cDNA clone was printed four times on the slide, the reproducibility of the replicated data could be judged using the coefficient of variation

(CV) of the Cy3 / Cy5 ratios of the replicated spots. The CV values for the mean, median and mode ratios of every clone on the array were calculated using the formula:

$$CV_{(\text{clone x})} = \frac{\text{Standard Deviation of replicated Cy3/Cy5}_{(\text{clone x})}}{\text{Mean of replicated Cy3/Cy5}_{(\text{clone x})}}$$

By averaging the CV for all cDNA clones, the global variation generated by using the mean, median or mode was calculated. The results are presented in table 3.1.

	MEAN	MEDIAN	MODE
Av. CV of replicated spot ratios	0.30	0.10	0.13

Table 3.1 The ratios for every spot from a Self – Self hybridisation on a custom microarray were calculated using the mean, median or mode of the total signal intensity. The CV values for each replicated clone were calculated and averaged to obtain the global CV.

The average CV of the replicated spot ratios using the mean is almost three times that obtained when the median is used, while the mode introduces similar variation to the median.

This variation for a single replicated clone is shown in figure 3.23. In figure 3.23a the mean, median or mode Cy3 intensities are plotted against the mean, median or mode Cy5 intensities respectively for every spot on the array. By focusing on one gene from the scatterplot (figure 3.23b) it is clear that the median measurement (red circles) introduced less variation amongst the replicated spot intensities. The mode measure appears to generate slightly more variability and the greater spread in the mean measurement of intensity suggests it is the poorest way to quantify fluorescence.

It has been shown (table 3.1) that the median is the best method for summarising the distribution of pixel intensities within a microarray spot and is therefore a more robust measure of central tendency. The large difference between the CV values obtained using the median and mean is possibly due to the fact that the median is less susceptible to outliers than the mean (Draghici S, 2003).

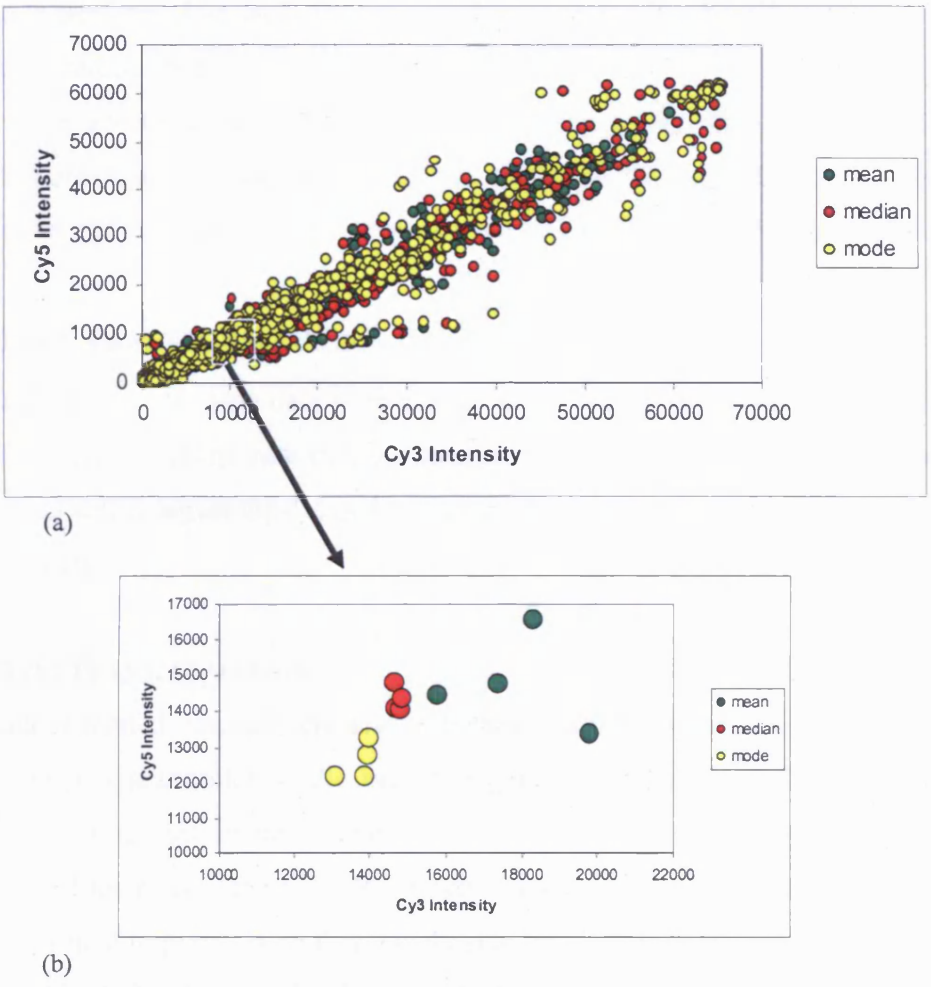


Figure 3.23 Presentation of Microarray Data

(a) A scatter plot of the Cy3 intensities vs the Cy5 intensities for spots on a Self vs Self hybridisation. For each spot on the custom array the mean (green), median (red) and mode (yellow) of the signal intensities are plotted.

(b) Focus on a single cDNA clone from within 3.27 (a). Each clone is printed four times on the chip. Imagene™ Software provided three measures of the signal intensity for each spot, the mean, median and mode. Note the tight clustering of the median values of the replicates compared to the greater spread in the data generated using the mean.

Outlying pixels in the spot sampling area can be caused by failings in the printing process which lead to uneven cDNA deposition or environmental contaminants such as dust particles. The process of segmentation (section 3.6.3) serves to remove these outliers but cannot be 100% efficient otherwise the mean and median ratios for the same spot would be equal. Nunez-Garcia et al., (2003) found the mode to be the most appropriate single-valued representative of the pixel intensity distribution within a microarray spot. Here,

the mode performs less well than the median at reducing the differences in intensities between replicate spots.

If the error generated by outlying pixels cannot be completely removed from the spot sampling area then it is imperative that their impact on the generated data is minimised ie the median of the signal intensity pixels should be used to quantify fluorescence.

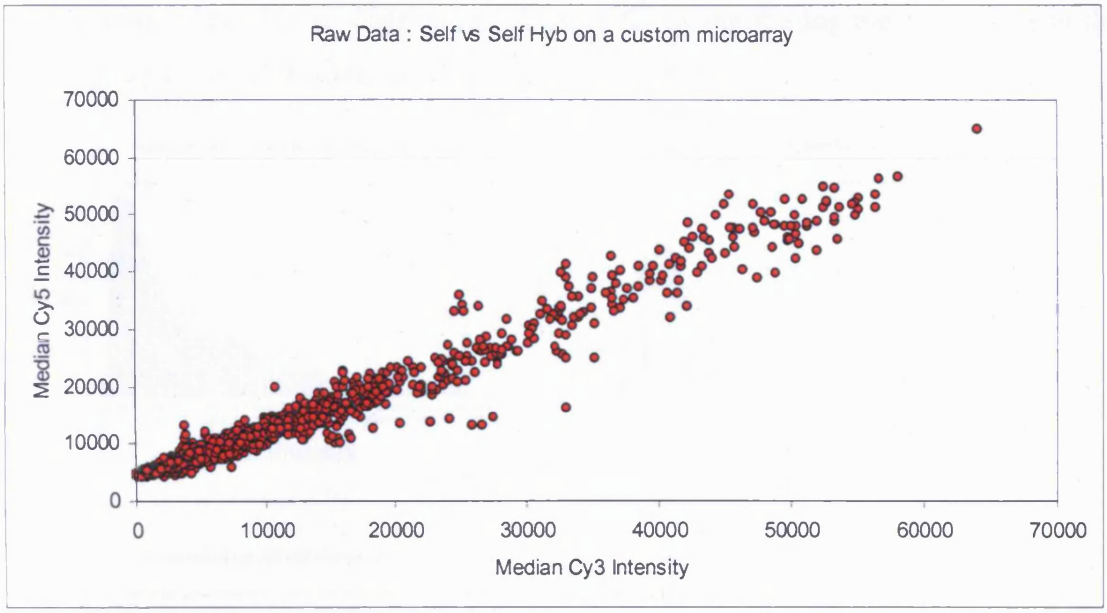
3.7 DATA TRANSFORMATION

The final step in analysing data from a microarray experiment involves the manipulation, or transformation of the raw data. Transformation is required to remove questionable data points and to adjust the data set to account for systematic errors that occur during the array process.

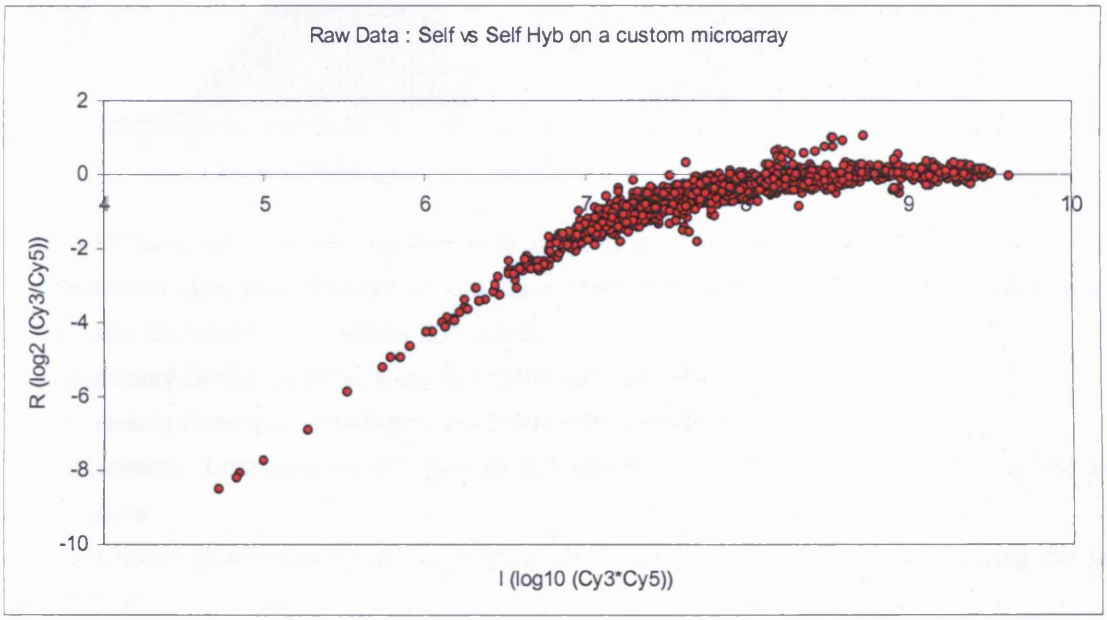
3.7.1 LOG TRANSFORMATION

The data generated from a microarray experiment can be viewed on a scatterplot. Figure 3.24a shows a scatterplot of the raw data generated from the Self- Self hybridisation carried out on a custom microarray slide. Each spot on the microarray provides two measures of intensity – the Cy3 fluorescence and the Cy5 fluorescence. The scatterplot gives a visual interpretation of this relationship by plotting Cy3 values against Cy5 values (Roberts CJ et al., 2000, ter Linde et al., 1999). As can be seen from figure 3.24a most points lie on the identity line suggesting, as expected, that the fluorescence at each spotted clone is roughly equal. Any differentially expressed genes will lie further away from this line. This means that the ratio Cy3/Cy5 will be approximately 1 for non-differentially expressed genes. A gene that is upregulated twofold in the Cy3 labelled sample will have a ratio of 2, whereas genes downregulated by the same amount will have a ratio of 0.5.

By taking the $\log_2$ (ratio) ie $\log_2$ (ratio) = $\log_2$ (Cy3 Intensity) - $\log_2$ (Cy5 Intensity), up and down regulated genes are treated similarly, such that a ratio of 2 becomes 1 and a ratio of 0.5 becomes -1. Non-differentially regulated genes will now have a ratio of 0. This simplifies the interpretation of data as the direction of differential expression becomes obvious and the magnitude of ratio changes are on the same scale (Yang YH et al., 2002). Another advantage of the log transformation is that it makes the distribution of the intensity values more symmetrical and closer to a normal distribution (Long et al., 2001,



(a)



(b)

Figure 3.24 a and b – Different ways of visualising the data generated by a microarray experiment.

(a) Scatterplot of the data from a Self vs Self hybridisation on a custom microarray chip. Each point represents the Cy3 intensity (x-axis) plotted against the Cy5 intensity (y-axis) for each spot on the microarray.

(b) An R I plot of the same data from a Self vs Self hybridisation on a custom microarray chip. Here the $\log_2$ (ratio) is plotted as a function of the $\log_{10}$ (Cy3 intensity * Cy5 intensity) for each spot on the array.

Yang YH et al., 2002). Figure 3.25 a and b show the distribution profiles before log transformation of the Cy3 and Cy5 intensities for all spots on the Self – Self

hybridisation. The skewed distribution changes following the log transformation to the bell shape of a normal distribution (figure 3.25 c and d).

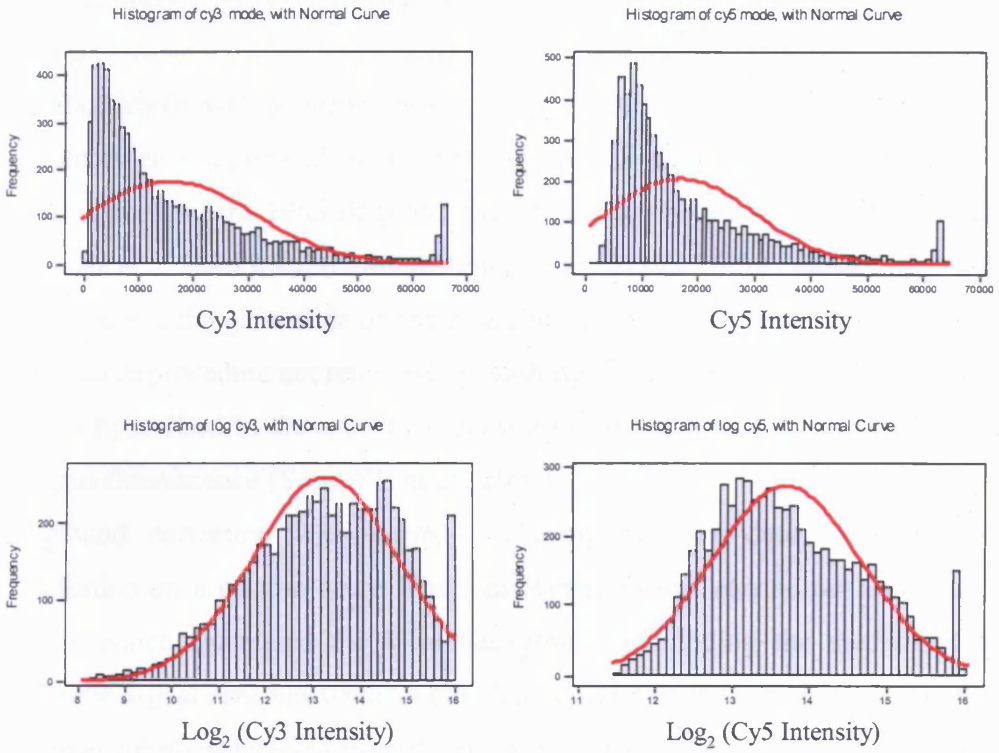


Figure 3.25 The effect of the $\log_2$ transformation on the distribution profiles of microarray intensity data. Note the skewed appearance of dataset towards lower intensities for the unlogged data. Log transformation tends to make the dataset more normally distributed.

(a) Cy3 intensity distribution for all spots on a 15000 Self – Self Hybridisation

(b) Cy5 intensity distribution for all spots on a 15000 Self – Self Hybridisation

(c) Cy3 intensity distribution for all spots on a 15000 Self – Self Hybridisation following the log transformation.

(d) Cy5 intensity distribution for all spots on a 15000 Self – Self Hybridisation following the log transformation.

Following the log transformation the data from a microarray experiment can be visualised on an RI (ratio vs intensity) plot (figure 3.24b). The RI plot displays the relationship between $\log_2$ (Cy3/Cy5) and $\log_{10}$ (Cy3*Cy5) and reveals systematic intensity dependent effects on the measured ratios (Quackenbush, 2002, Yang YH et al., 2002, Yang IV et al., 2002) not visible on a scatterplot. The data plotted in figure 3.24b is from the Self –Self hybridisation on the custom chip. In theory all points should lie on the line R=0 (ratio is 1) however at lower intensities the data deviates from the expected ratios and there is a considerable bias towards ratios with negative values. This suggests that at lower values

the Cy5 intensities for all spots are higher than those for the Cy3 intensities. This is possibly due to the additive effect of background fluorescence in the Cy5 channel or some systematic bias of the intensity on ratio.

3.7.2 BACKGROUND FLUORESCENCE

The fluorescence generated at a single microarray spot is thought to result from the additive effects of the labelled probe hybridising to the printed cDNA and background fluorescence. Background fluorescence may result from either (or both) auto-fluorescence of the glass slide or chemicals present in the various buffers used during the hybridisation procedure not removed by washing. Thus to quantify accurately the amount of probe hybridised to the spot it is necessary to subtract background fluorescence from the signal fluorescence (Yang YH et al., 2001).

Background correction was carried out using the raw data from the Self –Self hybridisation on a custom array. The transformation was carried out using GeneSight® software which calculated the local background by finding the median of all pixels outside the signal area but within individual squares of the sub grid (figure 3.21). This figure was simply subtracted from the signal intensity.

Compared with the raw data plot the background corrected data has lost the heavy bias towards spots with high negative ratios (figure 3.27a and b) suggesting that background was contributing to the observed ratios.

3.7.3 DETERMINATION OF A SPOT INTENSITY THRESHOLD

Whilst the RI plot for the background corrected data appears more symmetrical than that for the raw data (figure 3.27a and b), low intensity spots display greater variability in their ratios than spots with higher intensities. At low intensities more spots deviate from the expected $\log_2$ (Cy3/Cy5) of 0. Therefore, a method was needed that would remove spots whose intensities were close to background levels and would have contributed to noise in the data set.

Using the data from the Self – Self hybridisation, two methods were trialled to identify spots significantly above background levels and remove low intensity spots that contribute to noise in the data.

The first method calculated the mean and standard deviation (sd) of the local background of all spots on the array. A cut-off value was calculated using 2 x sd + mean (local

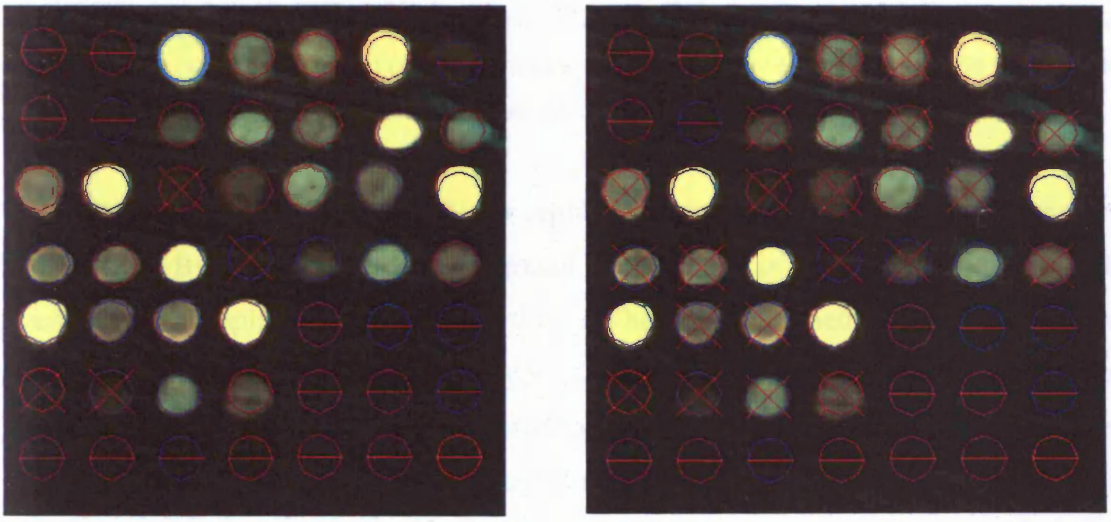
background of all spots). Spots were eliminated from the data set if they fell below this cut-off (Quackenbush, 2002).

In the second method the signal to noise ratios (SNR) for all spots were calculated by using the formula:

$$SNR = \frac{\text{(mean spot signal intensity-mean spot background intesity)}}{\text{Standard deviation Background intensity}}$$

The SNR is used in many signal-processing disciplines and is an indication of the confidence with which a signal can be quantified above background noise. Spots with an SNR less than 3 were excluded from further analysis.

The number of data points removed by each method can be seen in figures 3.27c and d. Using an SNR of three as a cut off point has removed far more data than the use of a value significantly greater than background. Using the background to calculate a cut off value seems a better approach than the SNR method as almost four times as much data is collected. Figure 3.26 shows a sub grid picked randomly from the custom array and the spots being excluded by each cut off method. The SNR method appears to remove spots that are well defined and clearly visible.



(a) (b)

Figure3.26 an example of spot exclusion using two different intensity based measures. Spots with red crosses are excluded from further analysis. Spots with lines are water blanks.

(a) Spots excluded due to signal intensity being less than two times standard deviation plus mean of the background intensity.

(b) Spots excluded due to signal to noise ratio being less than 3.

Before deciding on the method used in thresholding the quality of the data obtained first needed to be analysed.

Effect of transformations on data from Self-Self Hybridisation on custom array					
	Raw Data	Background Subtraction	Low spots removed using local background	Low spots removed using SNR	Normalisation
Mean $\log_2$ (Cy3/Cy5)	-0.63	0.19	0.20	0.17	0.07
Average CV for replicated spot ratios	0.12	0.17	0.03	0.5	0.04
Global CV	0.38	1.32	0.21	0.11	0.22
Number of data points	1536	1515	845	239	845

Table 3.2 The effect of various transformations on the data generated from a Self-Self hybridisation on a custom microarray.

The effects of the various transformations on microarray data was investigated. The aim of each transformation was to lower the average CV for replicated spot ratios in a self-self hybridisation to a custom microarray. As can be seen (row 4) removal of background significantly reduces variation in the data, and is further reduced by normalisation which removes systematic errors.

Both methods lower the average CV for replicated spot ratios compared to the raw data (table 3.2). By using the local background of all spots as an intensity threshold the average CV for replicated spot ratios is 0.03. When the SNR method is used the CV of the remaining replicates increases to 0.05. Thus, the local background cut off method introduces less variation amongst replicated spots than the SNR method despite removing four times the number of spots. Where the SNR method does perform better is in lowering the global CV. This decreased spread in data can be observed on the RI plots in figures 3.27c and d. This data spread is not necessarily the result of low intensity, or noisy measurements confounding real values. There is evidence that some genes preferentially incorporate one or other of the Cy dyes (Churchill, 2002). Therefore variations in the average ratios from 0, among spots, can occur by systematic errors such as preferential dye labelling. The global CV measure is not the best indicator as to which method should be used to determine a spot intensity threshold. In choosing a method the reproducibility of the replicated spots should be used. This means that the local

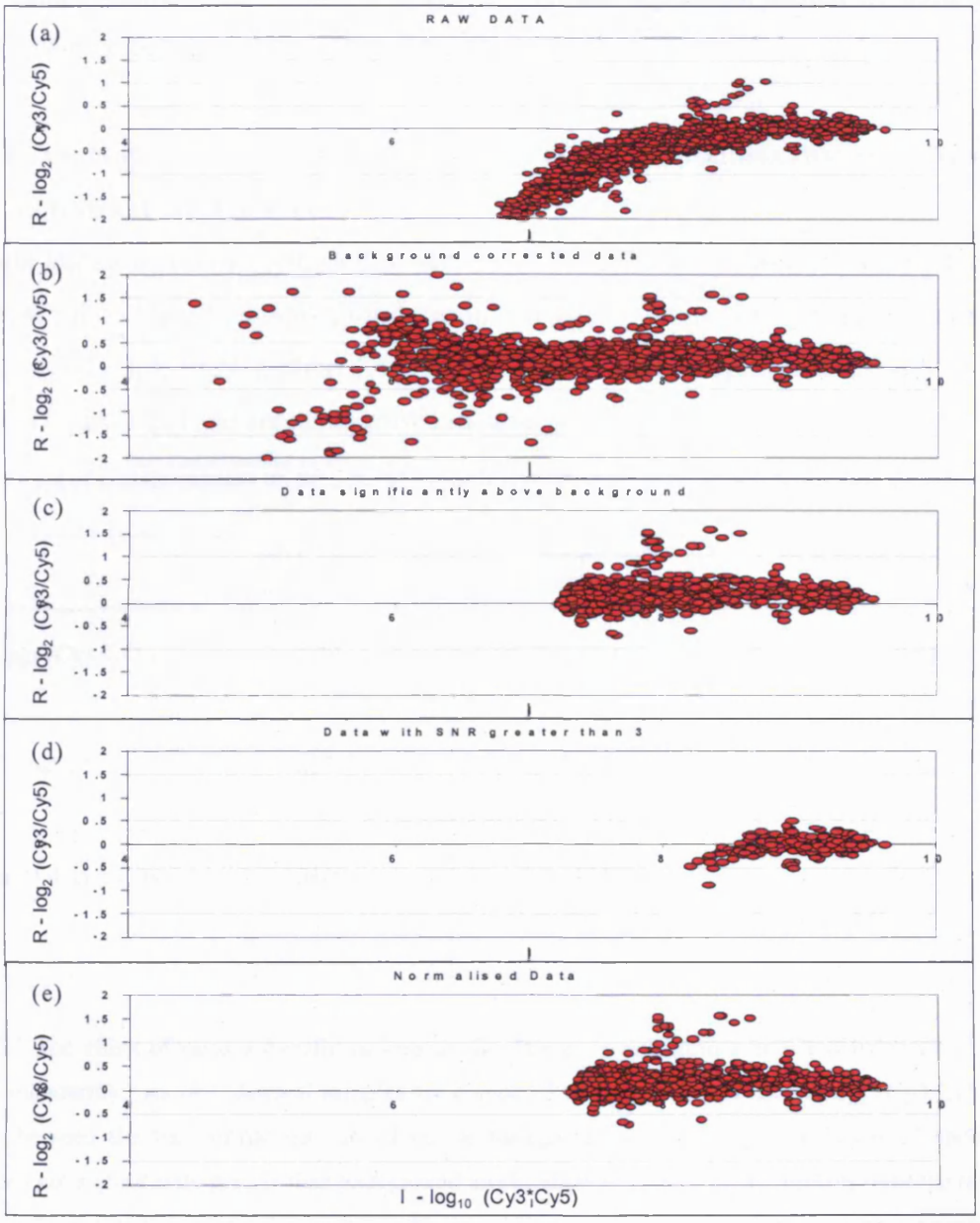


Figure 3.27 The effect of various transformations on microarray data – as revealed by RI plots.

RI plots are typically used to reveal systematic dependencies of the $\log_2$ (Cy3 intensity /Cy5 intensity) ratio on intensity. The $\log_2$ (Cy3 intensity /Cy5 intensity) or R, y-axis, is plotted as a function of the $\log_{10}$ (Cy3 intensity * Cy5 intensity), I, x-axis, for each spot on the array.

(a) RI plot of the raw data generated from a Self-Self Hybridisation to a custom microarray. (b) The RI plot is re-drawn after the local background fluorescence is subtracted from each spots raw fluorescence.

(c) Spots with signal intensites lower than two times the standard deviation plus the mean of the background fluorescence are removed from the analysis. (d) Spots with signal intensities that have a signal to noise ratio less than three are removed. Note fewer spots are evident indicating this method is more stringent than (c). (e) The data is normalised, following transformation (c), to correct for systematic errors using the mean of a group of invariant (housekeeping) genes.

background method, with a lower average CV for the replicates, should be used as the cut-off point for the removal of noisy data.

3.7.4 TRANSFORMATIONS OF DATA FROM A SELF-SELF HYBRIDISATION ON A NIA 15K CLONE SET MICROARRAY

The data transformations optimised in sections 3.7.1-3.7.3 were used to correct the data generated following a Self-Self hybridisation using labelled aRNA hybridised to an NIA 15K clone set microarray (section 3.4). The effect of the transformations are shown using RI plots (figure 3.28) and are summarised in table 3.3.

Effect of transformations on data from Self-Self Hybridisation on a 15000 NIA clone set array			
	Raw Data	Background Subtraction	Low spots removed using local background
Mean $\log_2$ (Cy3/Cy5)	-2.29	-1.6	-0.44
Global CV	1.33	4.44	0.42
Number of data points	15489	15489	7567

Table 3.3 The effect of various transformations on the data generated from a Self-Self hybridisation on a custom microarray. As two identical samples are compared each spot should have a $\log_2$ (Cy3/Cy5) of 0. As can be seen the transformations carried out ie background subtraction and removal of spots with intensities not significantly greater than background levels, show succesive improvements over the raw data in terms of bringing the average $\log_2$ (Cy3/Cy5) across the array closer to 0. Similar improvements, only following both transformations are observed for the global CV. Low CV values are reflective of data with low variance. Therefore removing low intensity spots has removed variability from the dataset.

From the data in table 3.3 it can be seen that the successive transformations are having the desired effect of bringing the mean log2 (Cy3/Cy5) of the data set closer to the expected value of 0. To analyse the spread in the data the global CV values must be used as the NIA 15K microarray does not contain any replicated spots. The effects of the transformations on the global CV are similar to those obtained when the data from the Self-Self hybridisation on the custom microarray was transformed. There is an initial increase in the global CV compared to the raw data following background correction, then a substantial decrease when spots with intensities below the background calculated

threshold are removed. Intensity thresholding has removed approximately 8000 spots, this data reduction is evident from the RI plots (figure 3.28c). Indeed the RI plots produced for each successive transformation for both the custom array data and the 15K clone set array data (figure 3.27 and 3.28 compared) are similar in shape.

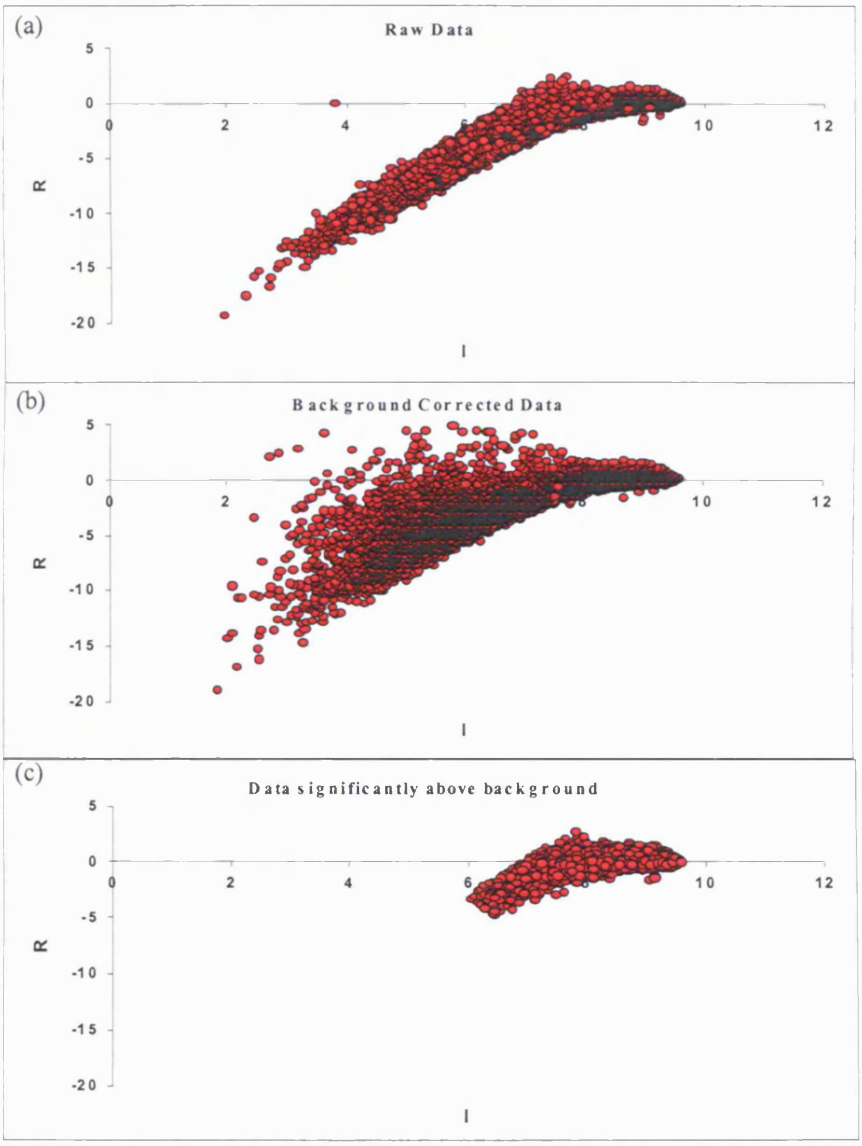


Figure 3.28 The effect of various transformations on the data from a Self-Self Hybridisation on an NIA 15K microarray. The $\log_2$ (Cy3 intensity /Cy5 intensity) or R, y-axis, is plotted as a function of the $\log_{10}$ (Cy3 intensity * Cy5 intensity), I, x-axis, for each spot on the array.

(a) RI plot of the raw data generated from the Self-Self Hybridisation

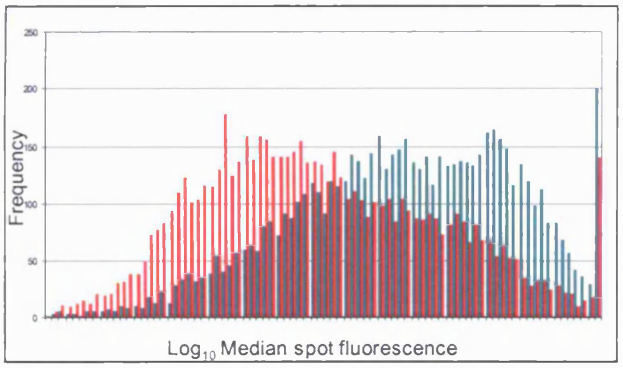
(b) The signal intensities for each spot are corrected by subtracting the local background fluorescence. The RI chart is re-plotted using the corrected data. Note the spread of the data at low intensities

(c) Spots with signal intensites lower than two times the standard deviation plus the mean are removed from the analysis. Thus fewer spots are visible at low intensities and the plot shows less variance in the data.

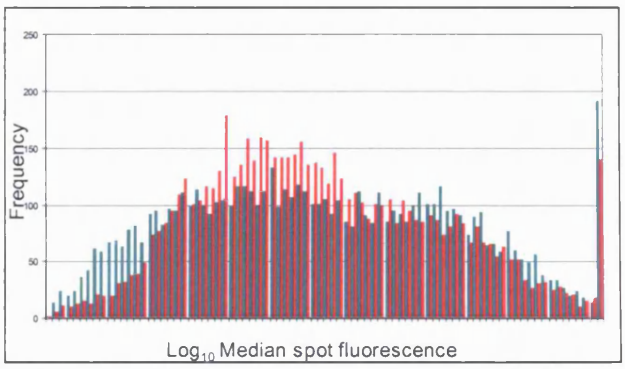
3.7.5 NORMALISATION OF NIA 15K MICROARRAY DATA

The purpose of normalising cDNA microarray data is to remove systematic biases in the measured expression levels (Quackenbush, 2002). GeneSight® Software provided various normalisations that could have been used to equalize the two channels. It was important that each normalisation method was tested to find the most efficient way of normalising cDNA data.

The Self-Self hybridisation provides a useful way of testing normalisations as the expected $\log_2$ (Cy3/Cy5) of every spot should be zero. Normalisations were tested on the semi-transformed data from the Self-Self hybridisation carried out on a 15000 NIA clone set (section 3.7.4).



(a)



(b)

Figure 3.29 The effect of a LOWESS normalisation on the distribution profiles of spot intensities in the Cy3 (green bars) and Cy5 (red bars) channels from a Self – Self hybridisation on a NIA 15,000 clone set array.

(a) Before normalisation both the Cy3 and Cy5 plots show the expected normal distribution, however the Cy3 intensity plot is shifted such that the mean of the Cy3 intensities is a higher value than the mean of the Cy5. This misalignment can be caused by a number of factors eg more RNA being labelled with Cy3 fluorophore, more efficient labelling in the Cy3 reaction.

(b) After a LOWESS normalisation both sets of data sit on top of each other suggesting that the correction has removed the bias in the data and the global intensities are equalised.

The underlying principle behind normalisation is that in a two-colour cDNA microarray hybridisation, the majority of genes will show no changes in expression between the RNA samples. This suggests that if equal amounts of RNA are labelled and hybridised to an array containing a random sampling of the organisms' genes, then the total fluorescence in the Cy3 channel and the Cy5 channel should be equal (Quackenbush, 2002).

In figure 3.29a histogram is shown for the distribution of the $\log_2$ Cy3 (green bars) and $\log_2$ Cy5 (red bars) intensities for all spots from a Self-Self hybridisation on a 15000 NIA clone set array. Although both channels show a normal distribution there is an obvious shift between the two data sets such that the Cy3 intensities are higher than Cy5 levels. There are a number of potential sources of bias inherent in the microarray process including unequal starting amounts of RNA being labelled with fluorophores, differences in the amount of labelled product generated by one or other of the fluorophores or discrepancies in the quantified fluorescence in either the Cy3 or Cy5 channel (Quackenbush, 2002, Leung and Cavalieri 2003). Any of these sources can lead to systematic errors in the data generated in a microarray experiment. Normalisation serves to equalize the mean intensity in the Cy3 channel and the mean intensity in the Cy5 channel. Following a LOWESS (for LOcally WEighted linear regrESSion) normalisation the data from the Self-Self hybridisation is re-plotted in figure 3.29b. The bias between the two channels has now been removed and the distribution plots are superimposed on one another.

The most basic normalisation involves equalizing the mean of the intensities of both the Cy3 and Cy5 channel (Quackenbush, 2002, Yang, YH et al., 2001b, Yang YH et al., 2002).

A normalisation factor (NF) is calculated by the formula:

$$NF = \frac{\text{Mean of Cy3 intensities}}{\text{Mean of Cy5 intensities}}$$

Every spot with a Cy5 intensity is then multiplied by the normalisation factor so that the corrected Cy3/Cy5 for each spot becomes:

$$R' = \frac{\text{Cy3 intensity}}{\text{NF* Cy5 intensity}}$$

The same principle is used during a median normalisation where the median intensity substitutes the mean intensities. A variation on the median normalisation is to equalize the intensities of the upper or lower quartiles of the intensity distributions. From table 3.4

it can be seen that normalising the semi transformed data using the mean, median or 0.25 percentile brings the average $\log_2$ ratio closer to 0. However by studying the RI plots of the data (figures 3.30 a – d) the limitations of these normalisations become apparent. For

Effect of Various Normalisation Strategies on data from Self-Self Hybridisation on a 15000 NIA clone set array							
	Low spots removed using local b/ground	mean	median	0.25 percentile	0.75 percentile	LOWESS	Piecewise
Mean $\log_2$ Cy3/Cy5	-0.44	0.18	-0.20	0.33	-0.53	-0.06	-0.11
Global CV	0.42	0.42	0.44	0.44	0.44	0.26	0.29
Number of data points	7567	7567	7567	7567	7567	7567	7567

Table 3.4 Effect of Various Normalisation Strategies on data from Self-Self Hybridisation on a 15000 NIA clone set array. The intention is to find the normalisation that gives a mean $\log_2$ Cy3/Cy5 closest to 0 (as it should be when no differential expression is occurring) and reduces the variation in the dataset as all $\log_2$ Cy3/Cy5 should be equal. Variation is calculated using the CV (std. deviation / mean). Thus, LOWESS appears to be the best normalisation method

the unnormalised data (figure 3.30a) the RI plot shows a clear deviation from ratios of 0 at low intensities. This "banana" shape is mirrored in the plots (figure 3.30b – e) for the normalised data. The global CV for the data following these normalisations changes very little from the unnormalised data suggesting that the normalisations have not lowered the standard deviation of the data. It can be concluded that, while these transformations

normalise spots with intensities close to the mean, median, 0.25 or 0.75 quartile the normalisations do not have the power to equalise spots over a wide range of intensities.
In contrast the LOWESS and piecewise normalisations deal with the increased variance in ratios at lower intensities. Indeed by studying the RI plots of the data following these normalisations (figures 3.30 f and g) it is clear that the "banana" shape has changed and most data points cluster on the line $\log_2$ ratio = 0. From table 3.4 the mean $\log_2$ ratio is -0.06 for the LOWESS and -0.11 for the piecewise normalisation. In addition the global CV has been reduced by almost a third (table 3.4) resulting in a much tighter data set.
The shapes of the RI plots in figures 3.30 a – e are typical of microarray data and reveal a systematic dependency of the $\log_2$ ratios on intensity. Both LOWESS and piecewise normalise microarray data in similar ways by removing this systematic dependency.
LOWESS uses a mathematical function that fits a smoothing curve to datasets that have non-uniform distributions, in this case the data in the RI plot. This is achieved by fitting a straight line through every point in the dataset. LOWESS calculates the gradient and intercept of the lines by using data points within a pre-defined window around the single data point. A polynomial function is fitted to the single data point using a least squares linear regression of the surrounding data points in the window. The linear regression is weighted such that the contribution of surrounding spots to the calculation diminishes with their distance from the single data point. The window slides along the data set and the procedure is carried out for every data point on the RI plot. When the entire dataset has been processed a smooth curve is produced providing a model for the data (Cleveland 1979, Cleveland and Devlin, 1988).
LOWESS was initially proposed as a normalisation for microarray data by Yang YH et al., (2002). For any point i on the RI plot the relationship between I (intensity) and R (ratio) is defined by $y(x_i)$, the LOWESS fit to the RI plot. Therefore, the LOWESS normalisation is carried out using the equation:

$$\log_2 (\text{Cy3 } i / \text{Cy5 } i)' = \log_2 (\text{Cy3 } i / \text{Cy5 } i) - y(x_i)$$

This equation is actually applied to the intensities before ratios are calculated such that:

$$\text{Cy3 } i' = \text{Cy3 } i \text{ and Cy5 } i' = \text{Cy5 } i * 2^{y(xi)}$$

where ' denotes the normalised values for a data point.

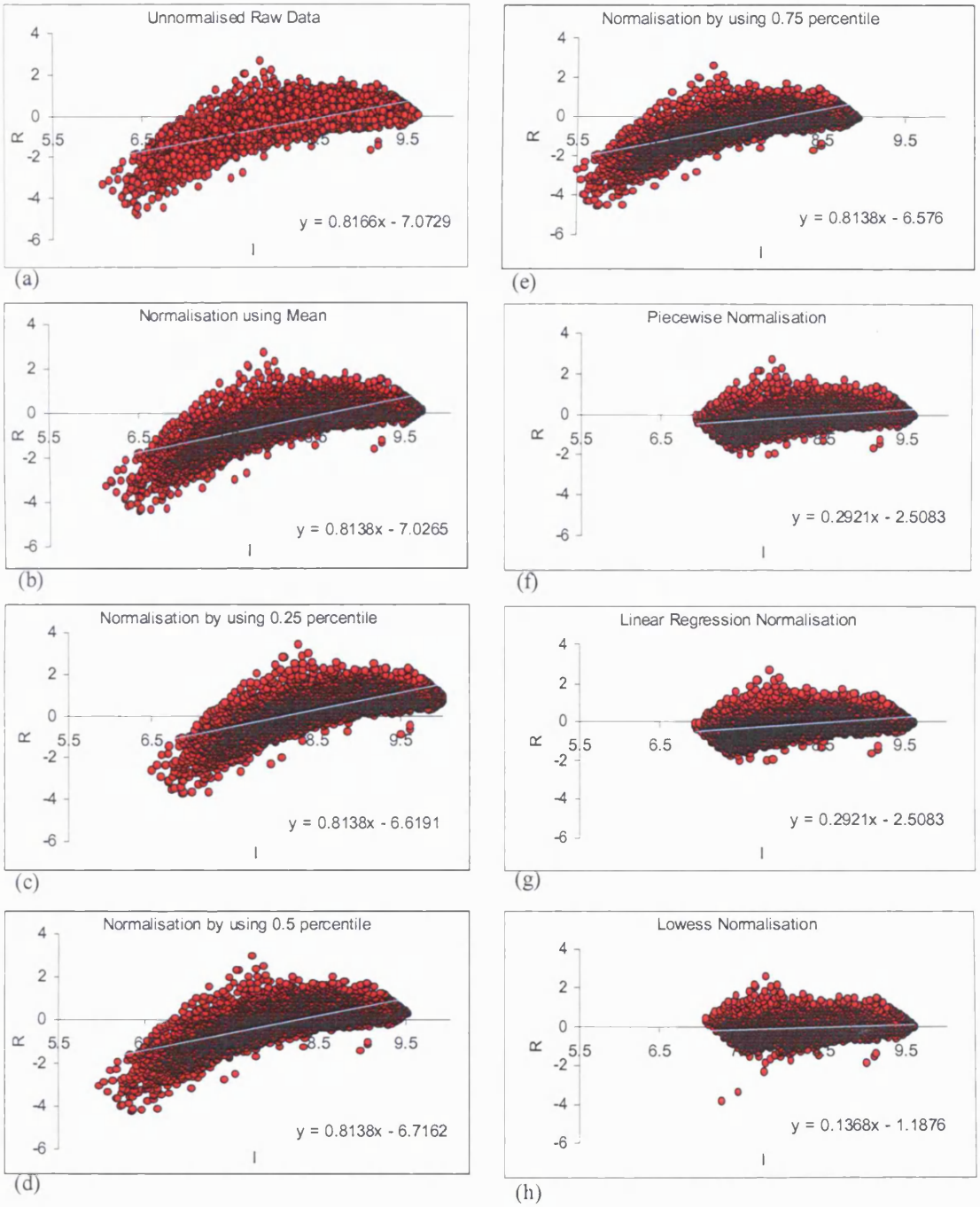


Figure 3.30 The effect of normalisation on the RI plots of the semi-transformed (background corrected and low intensity spots removed) data from a Self-Self hybridisation on a 15000 NIA clone set array.

(a) Un-normalised raw data

(b) – (h) The effect of various normalisation methods carried out using GeneSight® Software. Of interest is the shapes of the plots a – e when compared with plots f – h. In a – e the data points show a large deflection from the expected $\log_2$ (Cy3/Cy5) of 0 at low intensites. This effect is removed by using the normalisations shown in plots f – h suggesting that systematic error has been removed.

Piecewise normalisation works in much the same way as a LOWESS normalisation. The main difference is that the function $y(x_i)$ is not calculated for each individual point using

the sliding window. The data set is broken down with a fixed set of overlapping windows. In each fixed interval the data are approximated using a linear function. The linear function is then used to correct the intensities of the data points in the respective window as for the LOWESS normalisation. The advantage of the piecewise normalisation is that, theoretically, it is less computationally intensive than the LOWESS. In practice, using GeneSight® Software, no significant difference in processing time is observed between either of the normalisations. From the RI plots (figures 3.30 g and h) and the effect on the average mean and global CV (table 3.4) LOWESS and piecewise offer much improvement over the total intensity normalisations. LOWESS was therefore chosen as the normalisation for all data from hybridisations carried out on the NIA clone set microarrays.

3.7.6 NORMALISATION OF SMALL-SCALE CUSTOM MICROARRAY DATA

The normalisations described above are not applicable to hybridisations carried out using the custom microarray. As stated the underlying assumption behind normalisation is that the expression levels of the vast majority of genes do not change in the RNA samples being compared. This assumption holds when 15000 random genes are spotted onto a microarray slide (Quackenbush, 2002). However on the custom microarray the genes out were not randomly chosen. These candidate genes were picked for the microarray on the basis that their function is somehow linked to the process of neuronal plasticity (see chapter 4 for further discussion). This means that theoretically, following vibrissal deprivation and the subsequent induction of plasticity, most genes could show differential expression when compared with control samples. Normalisation would actually remove the differential gene expression negating the results of the experiment. To counter this problem the custom chip was designed to include housekeeping genes. These genes are constitutively expressed in all cell types and their expression is expected to be constant in control and experimental mRNA populations (Suzuki et al., 2000). Four housekeeping genes were included on the custom microarray; glyceraldehyde 3-phosphate dehydrogenase (GAPDH), β-Actin, Peptidylprolyl isomerase A or Cyclophilin A (PPIA) and Hypoxanthine Guanine Phosphoribosyl Transferase 1 (HPRT).

Each housekeeping gene had four replicated ratios which were first averaged. The custom microarray data was normalised by averaging the Cy3/Cy5 ratios of the housekeeping genes to obtain a normalisation factor. The Cy5 intensities of all spots were then corrected by multiplying by the normalisation factor. The effect of the normalisation on

the Self-Self hybridisation on the custom microarray is shown in figure 3.27e and summarised in table 3.2. Compared to the pre-normalised, semi-transformed data the normalisation has brought the average ratio of the data set closer to the expected value of 0 but has had little effect on either the average CV of the replicate spots or the global CV. By using a group of housekeeping genes that span the range of intensity levels of the scanner (figure 3.31), normalisation using this subset will not be biased towards genes in one intensity interval. Therefore the effect of using a total intensity normalisation, where the normalisation only seems to apply to genes near certain intensity levels eg the mean or median should not occur. Further, it has been shown that some housekeeping genes are differentially expressed in certain instances (Bustin, 2000, Dheda et al, 2004, Gibbs et al., 2003) therefore using more than one gene is recommended.

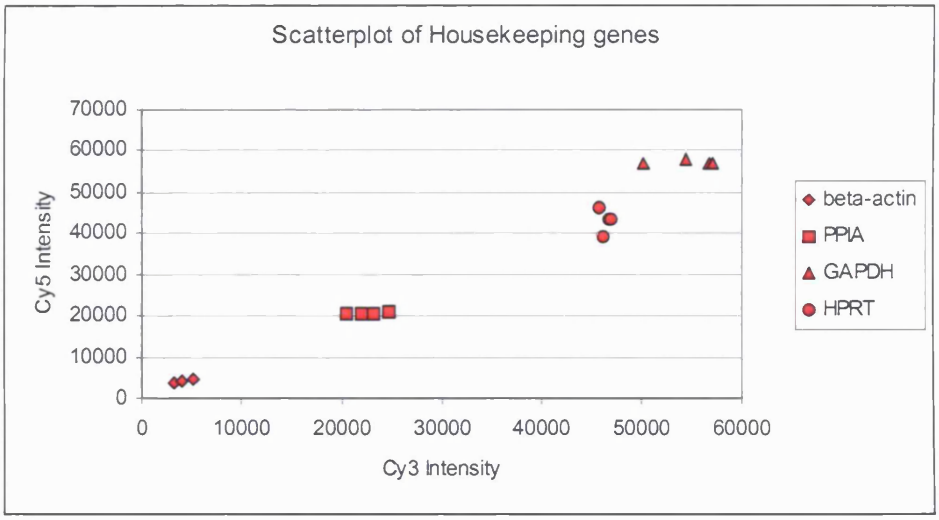


Figure 3.31 A scatterplot of the Cy3 intensity plotted against Cy5 intensity for the housekeeping genes present on the custom microarray. The data is taken form a Self-Self hybridisation and shows the range of intensities covered by the genes. By covering a wide range of intensities normalisation takes into account the systematic dependency of the ratio on the intensity inherent in cDNA microarrays.

3.8 STATISTICAL IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENES

In two colour microarrays two RNA populations, eg control and treatment are hybridised to a microarray slide. As each sample is labelled with a different fluorescent dye the relative expression levels of a particular gene in each sample is determined by the fluorescence generated by each dye at a particular spot (gene) on the microarray. For

each spot the relative expression levels of each dye are combined into a ratio providing a contrast of that particular gene's expression between the two samples.

Early microarray work identified differentially expressed genes by using an arbitrary cut-off value on the intensity ratios (typically 2 and 0.5, or 1 and -1 for $\log_2$ (ratio), Schena et al., 1996, DeRisi et al., 1997). This approach has a number of failings; the identification is not a statistical test, consequently, results have no confidence levels and the procedure is sensitive to variance heterogeneity across genes eg low intensity level genes tend to be classed as differentially expressed as microarray data shows significant variance at low intensity levels (Rocke and Durbin, 2001).

A number of methods have been proposed as alternatives to the fold cut-off and each method aims to:

a) reduce the number of false positives by taking into account the ratios of other genes on the array with similar intensities

b) reduce the numbers of false negatives by finding genes that are significantly differentially expressed at lower ratios.

Yang et al., proposed using the mean and standard deviation of genes within a sliding window of fixed width, and calculating an intensity dependent z-score for every gene on an RI plot. Genes that are greater than two standard deviations from their local mean ($Z>2$) are considered differentially expressed (Yang IV et al., 2002). Other methods involve selecting genes for which the ratio of the experiment and control values is a defined distance from the average ratio across the array (Tao et al., 1999).

In Chapter 4 differentially expressed genes will be identified using the Confidence Analysis tool within GeneSight® software suite (BioDiscovery, Inc.,CA). The function utilises the multiple spotting of genes on the custom microarray, allowing the inherent variability in measurements on a slide to be estimated (Kerr et al., 2000). Boundaries are established that separate ratio levels that are likely to occur from this inherent variability from those that occur due to true differential expression. This can be thought of as a noise sampling method (Draghici et al., 2003), that uses a log-linear statistical model and an ANOVA approach to model the noise characteristic of a particular microarray hybridisation, based on the following equation;

$$\log_2 R(gs) = \mu + G(g) + \epsilon(g,s)$$

where $\log_2 R(gs)$ is the measured log ratio for gene g and spot s

μ is the average log ratio over the whole array

$G(g)$ is a term for the differential regulation gene g

$\epsilon(g,s)$ is a zero mean noise term

an estimate μ' of the average log ratio μ is calculated by;

$\mu' = 1/n.m \, \Sigma_{g,s} \, log_2 (R(g,s))$

which is the sum of the log ratios for all genes and all spots divided by the total number of spots (m replicates and n genes).

An estimate ($G(g)'$) of the effect of gene g can be calculated as;

$G(g)' = 1/m \, \Sigma_g \, log_2 (R(g,s)) - \mu'$

which is the average log ratio of the spots corresponding to a given gene.

Using these estimates an estimation of the noise in a dataset can be calculated as;

$\epsilon(g,s)' = log_2 (R(g,s)) - \mu' - G(g)'$

Thus, for each spot on the array the noise is calculated as the difference between the actual measured $log_2 (R(g,s))$ and the estimated array effect μ' plus the gene effect $G(g)'$. This provides a noise sample for each spot and all the spots measured form an empirical noise distribution with a mean and deviation with associated confidence levels. The confidence intervals on the noise distribution can be mapped onto confidence intervals on the distribution of the ratios by bootstrapping (Draghici et al., 2003). The Confidence Analysis tool constructs several models that span the intensity range of the data so that low expressed genes use slightly different boundaries for the confidence limits.

One important feature of this test is that it does not require replication of hybridisations (although this is recommended) rather the confidence limits and noise distribution are calculated from replicated spots on a single array. This means the technique is useful in Chapter 4 where each gene is printed four times on a custom microarray and the amount of RNA available limits the number of possible replications. In practice 95% confidence intervals were chosen at an up/down regulation level of $\log_2$ (ratio) + or – 0.5.

In Chapter 5 through adequate biological replication statistical testing can be applied at the biological level. However the experiments will be conducted on NIA15K microarrays meaning the application of traditional statistical techniques is hampered by the sheer volume of data generated resulting in the problem of multiple testing. To illustrate the

problem of multiple testing a hypothetical experiment can be described in which it is required to identify, through microarray analysis, significantly differentially expressed genes between treatment and control RNA samples. A microarray hybridisation is carried out in which samples are co-hybridised to a microarray printed with 15,000 singly spotted cDNA targets. The experiment is replicated 5 times with independent biological replicates.

In order to identify genes significantly regulated between the samples, assuming the data is normally distributed, a t-test could be conducted and genes with p-values lower than a pre-specified cut-off, say 0.01, are declared to be significant. The p-value is the probability of rejecting a true null-hypothesis ie the probability of identifying a false positive. In this example, the cut-off significance level allows a level of confidence to be assigned to the list of regulated genes, that allow it to be stated that, 99 times out of a hundred, the test will correctly identify significantly regulated genes. Contrary to this, for every hundred times the test is run, a false positive will be declared (or a Type I error will be made). In the hypothetical example, if a t-test is conducted on each of the 15,000 genes, with a p-value less than 0.01 considered significant, then 1% of the tests will make a Type I error ie 150 genes will be declared significantly regulated between the two samples when in fact they are simply random effects. Thus, the probability of a false positive that is negligible if a single test is carried out is amplified when 15,000 tests are carried out in parallel (Draghici, 2003).

The problem of multiple testing is not new and numerous methods have been developed that attempt to control the so-called, Family-Wise Error Rate (FWER, Šidàk, 1967, Bonferroni, (Draghici, 2003)). However these corrections are overly restrictive and suggest that if, for example, 10,000 comparisons are carried out, a p-value less than$5x10^{-6}$ should be considered as siginificant. Such precision is beyond microarray technology and the applications of these corrections runs the risk of ruling out truly differentially expressed genes (false negatives) ie Type II errors will be made. Furthermore, both the Šidàk and Bonferroni corrections, in addition to the less conservative Holm's step down method, assume the variables being tested, in this case genes, are independent. This assumption rules out a known fact that networks of genes act in a co-ordinated, regulated manner (D'haeseleer et al., 2000). If the variables do act in a co-ordinated manner, then a modified version of Holm's step-down, the False Discovery Rate (FDR), method may be employed (Benjamini and Yekutieli, 2001).

This procedure corrects for multiple testing by essentially ordering genes in a list according to increasing p-values. The p-values of each gene, are compared to a threshold determined by the position of the gene in the ordered list. The threshold for the first gene in the list is determined by $(1/R).\alpha_e$ where R is the number of genes tested and α_e is the chosen significance level, in this example 0.01. The threshold for the k^{th} gene is $K/R.\alpha_e$. If the p-value for each gene is lower than the threshold then the null hypothesis can be rejected and the gene can be classed as significantly differentially expressed (Benjamini and Yekutieli, 2001).

The FDR method is elaborated in the Westfall-Young approach to multiple testing (Westfall and Young, 1993). Here, the labels applied to each group are randomly swapped around, whilst the data remains static. For each permutation of the labels the appropriate test statistic (eg t-test) is calculated, followed by Holm's step-down correction. This process is repeated thousands of times and a corrected p-value for each gene is the proportion of times the value of the test statistic calculated for the original data, is less than or equal to the value of the test statistic of a random permutation divided by the total number of permutations. This technique has been applied to microarray data and is a refinement of a procedure known as bootstrapping (Kerr and Churchill, 2001).

The Significance Analysis of Microarrays (SAM) software package was developed by Tusher et al. in response to the observation that the Westfall-Young procedure was still too conservative with their data (Tusher et al., 2001). SAM was therefore created to deal with this problem and is widely regarded in the microarray community as an efficient method for statistically analysing microarray data. This is ilustrated by the fact that Tusher et al.'s original publication describing SAM has, to date, been cited over 150 times, with the majority of these studies using SAM directly for analysis of microarray data (eg Masayesva et al., 2004).

SAM identifies statistically significant regulated genes by calculating gene specific test statistics similar to a t-test. This d-statistic is calculated as shown below:

$$d_i = \frac{\overline{x}_{i1} - \overline{x}_{i2}}{s_i + s_o} \qquad \textit{equation 5.1}$$

where: $\overline{x}_{i1}$ and $\overline{x}_{i2}$ are the average expression levels (in our case $\log_2$ (Exp/Ref), for treatments 1 and 2 respectively, and s_i is the standard deviation of repeated mesurements (an estimate of the population standard deviation). S_o is a so-called 'fudge-factor' which is added to prevent the test statistic d_i becoming too large when genes have variances close to 0 ie at very high or low expression levels.

Genes are then ranked according to the size of the calculated d_i . Permutation of the labels is carried out and for each permutation the d_i for every gene is calculated and ranked accordingly. Following the permutations the median (or 90th percentile) expected order statistic is estimated d_i* ie the median (or 90th percentile) d_i of the random permutations. A graph is then plotted of d_i versus d_i*. For the majority of genes, $d_i \approx d_i$*, but in a few cases, $d_i \approx d_i$* does not hold and these are genes that may be significantly regulated. A threshold, Δ, is applied, that is used to identify genes where $d_i - d_i^* > \Delta$. In the case of a two-sample comparison, thresholds are applied above and below the plot of d_i versus d_i* so that genes can be classed as up-, or down-regulated.

The size of Δ is changed and similarly up- / down-regulated genes are identified using these new thresholds. SAM then uses the permutated datasets and calculates the median and 90th percentile of the number of genes of all permutations, where $d_i - d_i^* > \Delta$ ie false positives. The median, or 90th percentile, of the number of falsely called genes is then divided by the number of genes called significant when the labels are correct (unpermuted), to give the FDR. The user then picks an appropriate Δ, preferably where the number of significant genes called is balanced with a low FDR (Tusher at., 2001). Depending on how stringent an investigator wishes to be Δ can be set low or high to retrieve a larger or smaller set of significant genes with a higher or lower FDR eg the experiment may be an initial pass to identify subsets of genes for further analysis. In this case a low Δ will generate more genes with a higher FDR. Thus the benefit of SAM is that typical cut-off p-values are not used to declare a gene as differentially expressed. Rather the significance of the result is generated using the data itself, therefore no assumptions are made regarding the distribution of the data. SAM is therefore different to traditional parametric statistical tests, which rely on prior knowledge about the distribution of the data.

3.9 TEST OF RNA AMPLIFICATION

RNA amplification is a well established technique that provides a potential solution to the problem that microarray experiments require a relatively large amount of total, or messenger RNA be used in probe production (Hegde et al., 2000, Van Gelder., et al 1990 and section 3.2.4). Indeed, as reported in section 3.2.4, a 1500 fold increase in the transcript population of barrel cortex total RNA has been obtained using Ambion®'s MessageAmp™ kit. It has also been shown that fluorescent probe can be generated from aRNA and in a self-self hybridisation, in which identical aliquots of aRNA are labelled with different fluorescent dyes and hybridised together, high quality data is obtained (sections 3.2.6 and 3.6.1).

PCR based methods are available for RNA amplification though it has been suggested that preferential amplification of certain transcripts tend to skew the expression data (Lockhart et al., 1996). Many studies have suggested that in-vitro transcription faithfully maintains the relative transcript levels in RNA samples, though none have directly tested the Ambion® MessageAmp™ kit (Baugh et al., 2001, Jenson et al., 2003). Prior to conducting microarray experiments the amplification linearity of this kit was tested.

An overview of the experimental design used to test the linearity of the Ambion® MessageAmp™ kit is shown in figure 3.32. All experimental samples were obtained from pooled total RNA (section 3.2.3) generated from mouse cortical tissue. Two aliquots of total RNA from this pool were converted to aRNA using the Ambion® MessageAmp™ kit as described (section 2.8.6). Both aliquots were suitably labelled and hybridised together on a microarray slide printed with the NIA 15K clone set (aRNA 1 v aRNA 2). Another aliquot of aRNA was co-hybridised with the total RNA from which it was generated (total RNA v aRNA 1). For a detailed overview of the experimental procedures used in these hybridisations see sections 2.16 – 2.19.

The results of the hybridisations are presented graphically in figure 3.33 and 3.34, in addition to the data obtained from a self–self hybridisation in which aliquots of the same aRNA sample were hybridised together (section 3.6.1). Shown on the scatterplots (figure 3.33 a-c), the R^2 value, or correlation co-efficient, for all three datasets is close to 1 indicating that the Cy3 and Cy5 intensity values of each experiment are tightly correlated and very few genes are differentially expressed.

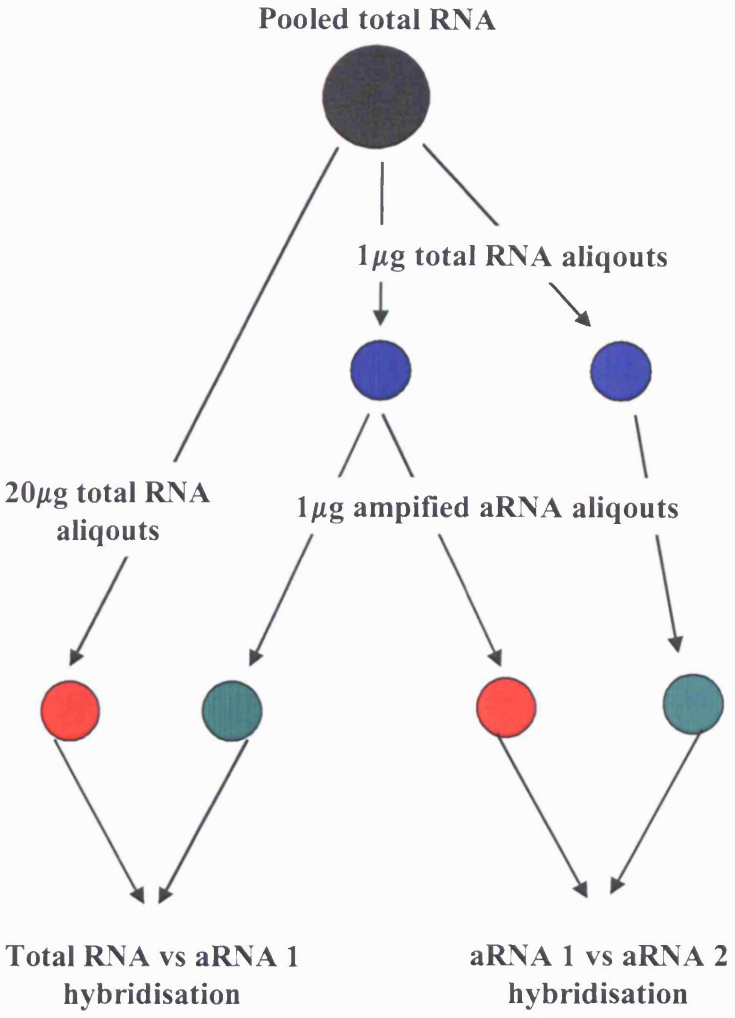


Figure 3.32 An overview of the experimental design used to test the linearity of total RNA amplification using Ambion® MessageAmp™ kit.

All samples used in the hybridisation where derived from a pool of total RNA (black circle) generated from mouse cortical tissue. Two 1μg aliquots (blue circles) were removed and amplified separately using Ambion® MessageAmp™ kit, to generate samples aRNA1 and aRNA2. In one hybridisation (aRNA 1 v aRNA 2), 1 μg of aRNA 1 was labelled with Cy5 (red circle) and 1μg of aRNA 2 was labelled with Cy3 (green circle), both were combined and hybridised to an NIA 15K clone set microarray.

In the second hybridisation, (total RNA v aRNA 1) 20μg of total RNA was labelled with Cy5 (red circle) and 1μg of aRNA 1 was labelled with Cy3 (green circle). Again both probes were combined and hybridised to an NIA 15K microarray.

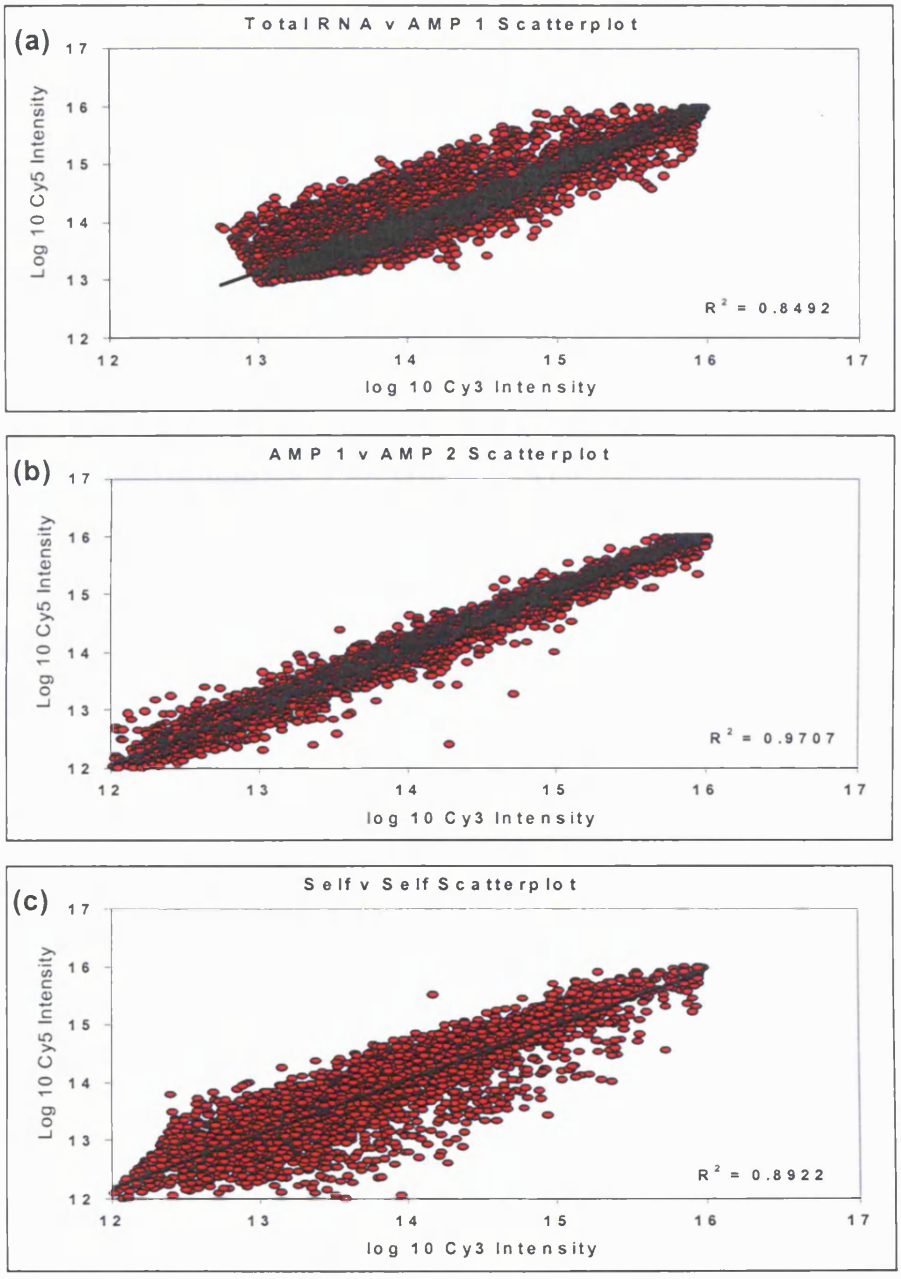


Figure 3.33 Scatterplots of data from three hybridisations used to test the linearity of RNA amplifcation using Ambion® MessageAmp™ kit. In each case the $\log_{10}$Cy3 intensity is plotted against the $\log_{10}$Cy5 Intensitiy for each spot. Provided on each plot is the correlation co-efficient (R^2), a measure of the similiarity of the Cy3 and Cy5 data for each microarray. From the self v self hybridisation (c) we can expect experimental variation to lower the expected R^2 from 1 (for perfectly correlated data) to 0.89. The R^2 values from the amplification test experiments, (a) and (b), show a high correlation in the data plotted with no significant difference to the self v self hybridisation. Thus it appears that RNA amplification does not distort the relative abundance of transcripts and the in – vitro transcription carried out using the Ambion® MessageAmp™ kit is indeed linear.

Interestingly, the R^2 value of AMP 1 v AMP 2 suggests that the data is more strongly correlated than the data from the Self-Self hybridisation (figures 3.33b and c) respectively). The least well correlated is the total RNA v AMP1 hybridisation data. This may be due to the fact that there is a large difference in the amount of labelled transcript in the two samples. If it is assumed that 1.5% of total RNA is messenger RNA (Alberts et al., 2002) then we are hybridising almost 3 times as much probe in AMP 1 which could affect the power with which the data is normalised. It could be the case, that normalisation works most effectively if approximately the same amount of probes are combined in a hybridisation (Quackenbush, 2002).

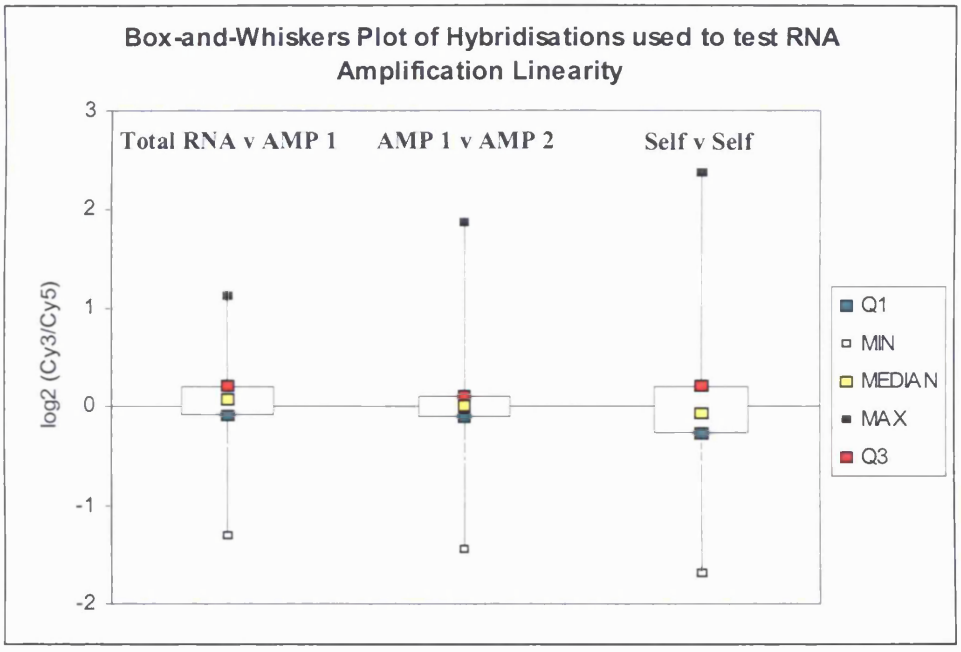


Figure 3.34 Box-and-Whiskers plots of the data for microarray experiments designed to test the linearity of RNA amplification using Ambion® MessageAmp™ kit. The plot contains five pieces of data that allows the distribution of the $\log_2$ (Cy3/Cy5) to be visualised for each microarray experiment. Within each box is a yellow square indicating the median of the distribution, at the edges of each box a green square indicates the lower quartile (Q1) of the distribution while a red square provides the position of the upper quartile (Q3). The whiskers projecting from the box end in a white (minimum $\log_2$ (Cy3/Cy5)) or black (maximum $\log_2$ (Cy3/Cy5)) square. The identity of each experiment is provided. From the plot it can be seen that the Self-Self hybridisation data shows the greatest variation in the ratios while the lowest variation is in the total RNA v AMP 1 data.

From the Box-and-Whiskers plot the greatest variation (the plot has longer whiskers and a wider box) is in the self-self hybridisation data (figure 3.34). This suggests that more genes show differential expression (a deviation from the expected ratio of 0) in this

hybridisation than those testing the amplification procedure. The point of including the self-self data was to provide an estimate of the random variation inherent in microarray data ie given that the log ratios of every gene will never equal 0, as they should in a self-self hybridisation, the data provides an estimate of how much the ratios will vary due to the impact of technical error. As both amplification test hybridisations, used total RNA from the same sample they can, in effect, be considered self-self hybridisations. If the in-vitro transcription was linear, then the amount of differential gene expression, or variation from the expected ratio (ie 0) in the amplification test hybridisations would be comparable to the results from a Self-Self hybridisation. This is indeed the case, in fact, the data from both amplification tests presents less variation than the direct self-self hybridisation data and is more reflective of the dataset expected from a self-self hybridisation.

Thus it can be concluded that the Ambion® MessageAmp™ kit does not introduce any significant non-linearity into the in-vitro transcription reaction and produces, on average, 40μg of aRNA from 1μg of total RNA.

3.10 DISCUSSION

This chapter began with an introduction to the methodology of microarray analysis (section 3.2). Prior to conducting experiments a literature search was carried out to identify the optimal conditions for generating an image and just as importantly, dealing with the large amounts of data generated. Through this research it was realised that there is no standard way of conducting a microarray experiment. At each step of the process debate rages as to the best methodologies to use; from the microarray platform used (Yauk et al., 2004), RNA labelling strategy (Hoen et al., 2003), quantification of scanned images (Yang YH et al., 2002) to the subjects of data handling, normalisation and statistical testing (Quackenbush, 2002, Churchill, 2002, Pan W, 2002) there appears to be no general consensus.

Therefore it was decided to optimise every step of the process, from generating an image, converting it to quantified data, to finally transforming the data, using the technology available at the time. Through this, standard operating procedures have been generated, that will save the time and expense of repeating sub-standard experiments and importantly, reduce potential errors that could have an accumulating effect on the data generated through this multi-step technology.

CHAPTER FOUR

CUSTOM MICROARRAY ANALYSIS OF GENE EXPRESSION IN THE BARREL CORTEX FOLLOWING VIBRISSAL DEPRIVATION

4.1 AIM

This chapter describes a study investigating the feasibility of microarray analysis in identifying genes, differentially expressed in the mouse barrel cortex, following vibrissal deprivation conditions that induce neuronal plasticity. To this end, a small-scale (custom) microarray was designed containing 371 genes. The majority of these genes were pre-selected on the basis that their encoded proteins have been either implicated in or may serve the purpose of facilitating the phenomenon of neuronal plasticity.

A thorough optimisation of every step in the microarray process was described in Chapter 3. Using this optimised work flow the effect of vibrissal deprivation, on gene expression in the mouse barrel cortex, will be investigated. A total of 371 genes, the majority of which are either known to be, or are hypothesised to be, functionally relevant to the process of neuronal plasticity, have been incorporated onto a microarray slide, which will be probed with total RNA probe derived from barrel cortices.

4.2 EXPERIMENTAL DESIGN

This section provides a detailed overview of the genes spotted onto the microarray, the pairing of samples on each microarray and the statistical procedures used to identify differentially expressed genes, prior to presenting results.

4.2.1 DESIGN OF A CUSTOM MICROARRAY

The 371 genes on the custom microarrays used in this chapter were chosen by conducting a comprehensive literature search primarily through the PubMed search tool at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed . A full list of the genes is provided in Appendix 1, column 1. Microarrays were generated using cDNA derived from bacterial clones of the required genes (section 2.15.1) obtained from three sources; the NIA 15K clone set (Tanaka et al., 2000), through the IMAGE consortium (distributed through Invitrogen, Paisley, UK) and through RT-PCR cloning procedures. Each clone from either the NIA, or IMAGE clone set has an individual ID, provided in column 3 of

Appendix 1, allowing the sequence of individual clones to be obtained, using sequence databanks such as Entrez Nucleotide;

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=nucleotide&cmd=search&term= .

The initial list from the literature search contained approximately 700 genes, of which only 300 could be identified in the NIA clone set, the most readily available source of clones. The time constraints imposed, by cloning genes from different sources, therefore reduced the number included on the microarray. There are a large number of physiological processes which could contribute to plasticity, each involving different genes. Therefore, while not all candidate plasticity genes could be included on the microarray, through outsourcing of clones it was ensured that at least some representatives of the most relevant physiological processes were included.

Columns 4, 5 and 6 (Appendix 1) provide links for each gene (where possible) to LocusLink, Unigene and Entrez Gene facilities within the NCBI website. These links will provide further information about the individual clone.

http://www.ncbi.nlm.nih.gov/LocusLink/

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=gene

The remainder of this section provides an overview of the genes on the microarray, by grouping them functionally. In column 7 of Appendix 1 is a keyword that assigns each gene to a particular group. Each keyword will be described, with the inclusion of examples of genes on the microarray assigned to each group, highlighting the relevance of each to neuronal plasticity. It should be noted that many genes are not restricted to one functional group and many genes show a degree of overlap between the groups.

Keyword CREB

Knowing that CRE-mediated gene expression accompanies certain whisker deprivation patterns that induce plasticity in the barrel cortex, genes were included on the microarray that play a role in in this transcriptional mechanism. At the centre of this group lies the transcription factor CREB. CREB is an inducible transcription factor that is activated when phosphorylated at Ser 133 and on activation binds to consensus (CRE) sequences within the promoters of genes. The phosphorylation of CREB allows it to dimerise and also to bind a transcriptional co-activator, CBP (CREB binding protein), which recruits components of the basal transcriptional machinery initiating gene expression. Along with CREM and ATF-1, CREB forms a sub-family of the bZIP superfamily of transcription factors (Lonze and Ginty, 2002, Mayr and Montminy, 2001).

CREB phosphorylation, and subsequent activation, is brought about by a wide range of physiological stimuli each capable of eliciting distinct cellular responses eg neuronal survival, neuronal proliferation, differentiation and axonal outgrowth in addition to learning, memory and plasticity. This suggests that the pattern of gene expression that results from CREB activation can vary dependent on the stimulus under investigation. Indeed recently published lists reveals a wide variety of genes known to contain CRE sequences whose functional proteins have widely differing physiological roles (Lonze and Ginty, 2002, Mayr and Montminy, 2001). The microarray was designed to include approximately 35 genes whose promoters contain the CRE sequence. Identified through a literature search, though not an exhaustive list, the list does include genes with a variety of functions, therefore the array is not limited to CREB target genes allied to a specific response to a specific stimuli.

The variations in the functional outcome of CREB activation are as diverse as (and possibly intrinsically linked to) the signalling pathways that lead to CREB activation. These pathways are described in more detail in section 1and although some cross-talk occurs between each they can broadly be classified into 4 paths:

1) cAMP signalling
2) Ca^{2+} signalling
3) Growth Factor Signalling
4) Stress-Induced signalling

Apart from stress-induced signalling, the pathways involve receptors on the post-synaptic membrane, which, on binding their respective ligands either undergo a conformational change which affects secondary messenger molecules or permit entry of Ca^{2+} into the cell which directly activates signal transduction effector molecules. Either way the signalling pathways involve the sequential phosphorylation / dephosphorylation of various kinases which leads ultimately to CREB phosphorylation and gene expression. Included on the microarray slide are some members of each signalling pathway in addition to the receptors and ligands that initiate certain cascades. Also genes that code for proteins involved in the CREB transcriptional complex ie *CBP* and *RNA polymerase II* are included.

Therefore including CREB family members, CREB target genes, and genes implicated in CREB activation signalling pathways on the microarray will provide important insights into the interactions between CREB family members, the functional significance of CRE-

mediated expression and the nature of the stimulus that leads to CREB activation within the mouse barrel cortex following whisker deprivation.

Keyword Calcium

In addition to the Ca^{2+} binding proteins that activate various signalling pathways that lead to CREB activation the microarray includes genes for proteins that form Ca^{2+} permeable ion channels, pumps and exchangers on the basis that vibrissal deprivation may direct the synthesis of proteins that actively regulate the intracellular free Ca^{2+} levels in neurons.

Keyword CPG

A number of paradigms are used to study the phenomenon of neuronal plasticity. Although the paradigms study plasticity in various organisms, or in different areas of the central nervous system of the same organism, a common assumption across all models is that gene transcription and subsequent *de novo* protein synthesis are necessary for, or accompany, the observed long term changes. For example gene transcription in the mouse hippocampus is a pre-requisite for LTP and is also necessary in the amygdala during fear conditioning, a form of associative learning thought to be reliant on plasticity (Frey et al., 1988, 1996, Bailey et al., 1999). Furthermore gene transcription is necessary for long term facilitation in the sea snail *Aplysia* (Castellucci et al., 1986) and is associated with long-term plasticity in the neuromuscular junction of the fruit fly *Drosophila* (Dubnau and Tully, 1998). Researchers have therefore sought to identify these genes and a host of candidate plasticity genes (CPG) have been published.

Many of these CPGs are not exclusive to one particular model of plasticity eg levels of the neurotrophic factor BDNF increase within the CA1 region of the hippocampus following LTP induction (Patterson et al., 1992), in the barrel cortex of the rat on unilateral vibrissae deprivation (Nanda and Mack, 2000) and BDNF enhances long term facilitation in *Aplysia* (Purcell et al., 2003). Another example of a family of genes with multi – model plasticity involvement are the *cadherins*, cell adhesion molecules that have been implicated in developmental plasticity of the rodent barrel field and visual systems in the rodent and drosophila (Huntley and Benson, 1999, Inoue and Sanes 1997, Lee et al., 2001) and are known to be necessary for the induction of LTP (Tang et al., 1998). In contrast the *activin beta A* subunit gene, a putative CPG, shows a marked increase in expression on induction of hippocampal LTP but has not been studied in other plasticity models (Inokuchi et al., 1996).

A list of CPGs for inclusion on the custom microarray has therefore been drawn from the wide variety of neuronal plasticity models currently studied in a diverse range of

organisms, on the basis that many genes are known to be functionally involved in a number of different paradigms. Attempts were made, in compiling a list, to encompass as many CPGs as possible, whether they are common to many models or have been identified in just one as theoretically any could be involved in experience dependent neuronal plasticity in the mouse barrel cortex.

Keyword Synaptogenesis

In comprising a list of genes for inclusion on a custom microarray a fundamental question needed to be asked. What cellular changes occur in the barrel cortex, on whisker deprivation, that facilitate the plastic changes observed? It can be hypothesised that new synapses are formed, or existing synapses are somehow modified, providing enhanced points of contact between neurons from adjacent barrels.

Both synaptogenesis and synaptic restructuring have been proposed as models to account for neuronal plasticity in many paradigms. For example, when LTP is initiated in a small area of CA1 of the hippocampus ($30\mu m^2$) newly formed dendrites are observed (Engert and Bonhoeffer, 1999) and are accompanied by an increase in multiple synaptic boutons (Toni et al., 1999) in a model which proposes that small synapses enlarge, become perforated and develop a complete spine partition and segmented post-synaptic density (PSD) via an intermediate synapse with sectional partition and horseshoe shaped PSD (Chang and Greenough, 1984, Desmond and Levy, 1986, 1990).

Numerous studies have reported increases in synaptic density occur within specific loci of the CNS as a result of behavioural learning experiences eg in the subfield CA1 of rats after brightness discrimination conditioning in a Y-maze (Wenzel et al., 1980); in the *robustus archistralis* nucleus of canaries following acquisition of male-like singing behavior by females treated with testosterone (DeVoogd et al., 1985); and in the motor and cerebellar cortices of rats after motor skill learning (Black et al., 1990; Kleim et al., 1996, 1997).

Synaptogenesis is directed by axonal growth cones which direct the developing axons to dendritic targets. Following cell-cell contact, cell-adhesion molecules stabilise these contact points and pre- and post-synaptic differentiation occurs to form a functioning synapse (Garner et al., 2002). The most widely studied cell-adhesion molecules are the cadherins (previously mentioned as CPGs) which are trans-membrane proteins capable of homophillic or heterophillic binding. Found pre- and post-synaptically (Benson and Tanaka, 1998) cadherins, in particular N-CAM, are thought to stabilize the site of cell-cell contact through trans-synaptic binding. Maintenance of the rudimentary and fully

functional synaptic connection is thought to be mediated by a complex formed between the intra-cellular N-CAM and α-catenin, a molecule which interacts with actin and actinin, a model that provides a link between the cytoskeleton and the synapse (Togashi et al., 2002, Murase et al., 2002). In addition the extracellular domains of members of the pre-synaptic neurexin protein family are able to stabilize cell-cell contacts via direct interactions with the extracellular domain of the post-synaptic protein neuroligin (Nguyen and Sudhof, 1997). Cytoplasmic domains of β-neurexin bind to CASK (Hata et al., 1996), a presynaptic scaffold protein, while similar cytoplasmic domains in neuroligin bind to the post-synaptic density (PSD) protein SAP90/PSD95 (Irie et al., 1997, Bolliger et al., 2001). CASK binding leads to the formation of a pre-synaptic macromolecular complex including Ca^{2+} channels (Butz et al., 1998, Maximov et al., 1999) while the neuroligin / SAP90/PSD95 complex is known to bind, amongst other molecules, subunits of NMDA receptors leading to the hypothesis that complexes formed by neurexins and neuroligin maintain pre-synaptic vesicular release and postsynaptic receptor functioning. Thus the genes that code for N-CAM, along with other members of the *cadherin* gene family, are included on the array in addition to *neurexin* and *neuroligin*.

Keyword Pre-Synaptic

In addition to genes involved in establishing initial cell–cell contacts, attempts were made to create a microarray containing a number of genes that code for proteins functionally relevant to the mechanics of the synapse.

Presynaptically, of particular interest are the genes for proteins involved in the exocytosis of synaptic vesicles which mediate neurotransmitter release. Neurotransmitter release is a cyclical process initiated by action potentials which trigger neurotransmitter release from the presynaptic membrane (Katz, 1969). Proton pumps in the synaptic vesicles are responsible for neurotransmitter uptake to form a pool of vesicle clusters in the presynaptic nerve terminal (Maycox et al., 1988). Members of the SNARE family of proteins, located on the intracellular membrane, dock and prime filled vesicles ready for neurotransmitter release (Jahn and Sudhof, 1999). A full list of SNARE proteins is published (Hay, 2001) and the microarray contains both presynaptic membrane bound SNAREs (*Syntaxin 1B* and *SNAP 25*) and one member of the family, *VAMP*, that binds to the synaptic vesicles themselves (Söllner et al., 1993). Also included is a member of the synaptophysin family which may play a role in regulating SNARE function by binding to VAMP (thus preventing VAMP from taking part in vesicular–membrane fusion (Calakos and Scheller, 1994). Vesicle fusion with the presynaptic membrane is thought to be

controlled by SM (Sec1/Munc18-like) proteins (included on the array, Jahn et al., 2003). Rapid exocytosis of the vesicle may be achieved through the actions of synaptotagmins, putative Ca^{2+} sensors that when they bind Ca^{2+} are proposed to undergo a conformational change and no longer bind the SNARE/vesicle complex instead interact with the membrane. This dissociation puts the SNARE/vesicle complex under mechanical stress forcing open the fusion pore (Geppert et al., 1994, Voets et al., 2001, Sŭdhof, 2004) resulting in neurotransmitter release into the synaptic cleft.

Components of synaptic vesicles included on the microarray are the *SV2* gene and members of the *synapsin* gene family (Buckley and Kelly, 1985, Greengard et al., 1994). SV2 may be a Ca^{2+} transporter in synaptic vesicles while the exact function of synapsins is unknown. It is known that synapsins is that when phosphorylated eg by CAMKI or PKA they no longer bind to vesicles (Hosaka et al., 1999). Also *synapsinI* knockout mice show a significant loss of synaptic vesicles suggesting a role in the maintenance of the vesicular pool (Rosahl., et al 1993). Futhermore deletion of *synapsinI* delays axonal outgrowth and synaptogenesis of cultured neurons (Ferreira et al., 1998) which emphasises their importance and their inclusion on the microarray.

Keyword Post-Synaptic

The inclusion of genes whose encoded proteins are involved post-synaptically is as important as the pre-synaptic genes.

At the site of synaptic contact the postsynaptic membrane thickens to form the postsynaptic density (PSD), a specialised organelle whose prime purpose is receiving and transducing signal from the pre-synaptic neuron to the post synaptic neuron. In the majority of excitatory synapses in the mammalian CNS this signal is the neurotransmitter glutamate and prime constituents of the PSD are the glutamate receptors of which the N-methyl-D-aspartate (NMDA), α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) and the group I metabotropic glutamate receptors (mGluRs) are specifically targeted to the postsynaptic membrane (Takumi et al., 1999).

NMDA receptors are heteromeric complexes composed of NR1 and NR2 (a to d) subunits that show similar membrane topology to the GluR 1-4 subunits that make up AMPA receptors (Dingledine et al., 1999, Hollmann and Heinemann, 1994). Seven mGluRs have been identified each containing 7 trans-membrane domains (Conn and Pin, 1997). All three receptor types have been implicated in neuronal plasticity and for this reason and their importance in the functioning of the synapse, all *NMDA* and *AMPA* subunit genes, in addition to members of the *mGluRs* have been included on the array.

In particular the NMDA receptors are central to the process of LTP (section 1.) and AMPA receptors, that are rapidly inserted into the postsynaptic membrane and mediate most rapid excitatory neurotransmission (Liao D et al., 1995, Isaac et al., 1995, Durand et al., 1996) are also required for induction and maintenance of LTP and for rapidly acquired spatial learning (Lee et al., 2003).

The lack of AMPA receptors at some excitatory synapses and the universal presence of NMDA receptors suggests that the two are targeted at synapses by different mechanisms (Nusser et al., 1998, Petralia et al., 1998). The proteins involved in targeting the glutamate receptors to the postsynaptic membrane also contribute to the PSD machinery and a number of genes encoding these proteins are included on the microarray.

The protein PSD-95 binds NMDA subunits through an interaction of its PDZ domain and the cytoplasmic C-terminus of the subunit (Kornau et al., 1997 a and b). This interaction is important for a number of reasons; it holds the receptors in specific locations in the PSD and couples receptors to signal transduction proteins (Steigerwald et al., 2000, Migaud et al., 1998, Sprengel et al., 1998). Examples of signal transducers included on the array are *SynGAP* which binds the PDZ domains of PSD-95, a member of the Shank family (which is bound to the PSD-95 via an intermediate protein), and Homer which interacts with Shank (Kim et al., 1998, Chen et al., 1998, Naisbitt et al., 1997, Naisbitt et al., 1999, Tu et al., 1999). Homer is a cytoplasmic adaptor protein which was initially discovered though its interactions with mGluRs and can conceivably link NMDA receptors with the mGluRs (Brakeman et al., 1997) and indirectly, through intermediates, with intracellular Ca^{2+} stores (Tu et al., 1999). It has been proposed that Homer and Shank may ineract with the postsynaptic skeleton to produce physiological changes eg enlargement of dendritic spines in response to synaptic neurotransmission (Sheng, 2001, Naisbitt et al., 1999).

The interactions of AMPA receptors with the PSD involves a different set of proteins, possibly reflective of the rapid regulation of AMPA at synapses (Shi et al.,1999, Hayashi et al., 2000). Two genes on the microarray, *Pick1* and *Grip*, bind AMPA receptors in a phosphorylation dependent manner. Blocking the GluR2-GRIP interaction is known to prevent synaptic potentiation, suggesting GRIP is involved with receptor recruitment (Li et al., 1999). AMPA receptor phosphorylation prevents GRIP, but not PICK, binding and this could provide a phosphorylation dependent switch that would regulate AMPA sorting (Xia et al., 1999, Chung et al., 2000, Matsuda et al 1999).

Keyword Cytoskeleton

The neuronal cytoskeleton is composed of microfilaments, intermediate filaments and microtubules and is a major regulator of neuronal morphology. During development co-ordinated changes occur in the actin and microtubule cytoskeleton that enable eg axonal outgrowth and dendrite formation and it can be hypothesised that cytoskeletal remodelling in the adult contributes to the process of neuronal plasticity, therefore key players involved in cytoskeletal organisation have been included on the array (Luo, 2002).

This group of genes include the basic elements of the cytoskeleton themselves eg *β-actin*, α- and *β-tubulin* and *neurofilament-L*, in addition to genes for proteins involved in its regulation. The dynamics of the actin cytoskeleton, at all stages of an organisms lifespan, are directed by signalling pathways that take their cues from receptors on the neurons membrane responding to extracellular signals (Luo, 2002). The best understood is the RhoGTPase pathway, important for a number of reasons:

a) they are key regulators of the actin cytoskeleton in many cell types (Van Aelst and D'Souza-Schorey, 1997)

b) RhoGTPases regulate myosin activity and actin polymerisation and play an important part in neuronal morphogenesis (Dickson, 2001, Luo, 2000, Redmond and Ghosh, 2001).

Key RhoGTPases included on the microarray are *Cdc42* which regulates filopodial outgrowth and *de novo* actin polymerisation (Kozma et al., 1995, Nobes and Hall, 1995), Rac which regulates lamellipodia (Ridley et al., 1992, Nobes and Hall, 1995) and RhoA which acts on the actin cytoskeleton indirectly via the regulation of myosin II. Myosin II is known to force actin fibers into bundles driving growth cones forward (Svitkina et al., 1997) and is positively regulated through phosphorylation by Rho-associated kinase/ROCK (Amano et al., 1996) which is the primary effector of RhoA (Kaibuchi et al., 1999) and is also included on the array. Microtubules serve as scaffolding for neurite outgrowth and facilitate vesicular transport. Built from α- and β-tubulin monomers these highly dynamic structures are regulated by microtubule-binding proteins (Mandelkow et al., 1993, Desai and Mitchison, 1997, Wiche et al., 1991). One sub-group, the microtubule-associated proteins (MAPs) have been implicated as primary targets of signalling pathways implicated in neuronal plasticity.

Invariably the phosphorylation of MAPs by kinases and phosphatases from the signalling pathways affects their binding to microtubules thus affecting the cytoskeleton and possibly facilitating plasticity (Drewes et al., 1998, Tokuda and Hatase, 1998, Gundersen and Cook, 1999). Included on the microarray is a variety of *MAPs* including *MAP2* and *MAP1B*. MAP2 stabilizes microtubules leading to prolonged growth periods and net accumulation of microtubules in cellular and neurite outgrowth (Drewes et al., 1998, Gamblin et al., 1996, Kowalski and Williams, 1993) and it is proposed that redundancy may occur with MAP1B, as double knockouts present defective neurite outgrowth (Teng et al., 2001, Takei et al., 2000). Both MAP2 and MAP1B are known to interact with the actin cytoskeleton and are clearly inportant in regulating neuromorphogenesis and neurite initiation (Pedrotti and Islam, 1996, Togel et al., 1998, Sattilaro, 1986).

The actin and microtubule cytoskeleton can also be linked by Plakins, one member Macf functions as a versatile cytoskeletal linker protein, playing a key role in neural development (Bernier et al., 2000).

Microtubules form long stiff bundles lending rigidity to the neuronal cytoskeleton and it has been proposed that MAP2 dimerization cross-links and stabilises these bundles (Wille et al 1992a and b, Chen et al., 1992, Leclerc et al., 1993). MAP phosphorylation may be the link between cytoskeletal changes in response to neuronal signalling eg MAP2 is a substrate for many protein kinases and phosphatases including PKA, CAMKII and PKC (Walaas and Nairn, 1989) proteins implicated in plasticity (Brandon et al., 1997, Glazewski et al., 1996, Pascale et al., 1998) and in CREB activation.

In addition genes encoding members of a large group of proteins, called MAPKs (mitogen-activated protein kinases) or ERKs (extracellular signal-regulated kinases), have been included on the basis of their ability to phosphorylate MAPs and possible participation in the regulation of synaptic transmission and long-term memory (Boulton et al., 1991, Cobb et al., 1991, Brambilla et al., 1997, Kornhauser and Greenberg, 1997). In addition phosphatases that catalyze the de-phosphorylation of MAPs, in particular MAP2, including PP2A and Calcineurin (Guerini, 1997, Sim, 1991, Goldberg, 1999) as these phosphatases have been shown to be essential for the regulation of neuronal plasticity and synaptic transmission (Sim, 1991, Nairn and Shenolikar, 1992).

The phosphorylation state of MAP2 alters the dynamics of microtubules by affecting its binding to them. Phosphorylation of MAP2 appears to reduce the interaction with microtubules and this is reversed when MAP2 is dephosphorylated, however the degree of phosphorylation appears to be less important than the specific site of phosphorylation

in determining the association (Tsuyama et al., 1987). Phosphorylation at different sites on MAP2 may influence different stages of microtubule assembly and these distinct sites appear to be phosphorylated by different kinases (Sánchez et al., 2000). Thus a possible scenario exists whereby, repeated synaptic activity results in an alteration of intracellular Ca^{2+} levels, leading to the activation of kinases/phosphatases which in turn leads to changes in MAP phosphorylation and a more dynamic cytoskeleton. This could potentially lead to changes in synapse morphology and/or the number of synaptic contacts therby leading to a potentiation or depression of synaptic transmission. Indeed MAP2 phosphorylation modifications have been found in several neuronal plasticity models including visual cortex plasticity (Reid and Daw, 1995, Aoki and Siekevitz, 1985) and in LTP (Dīaz-Nido et al., 1993, Quinlan and Halpain 1996a).

Keyword Glutamate

A group of genes have been included that act in both the synthesis and reuptake of glutamate the primary neurotransmitter of excitatory synapses in the CNS. These include *Glud1*, *Glast-1* and *Got1* (Levenson et al., 2002). These genes are included on the basis that inceased glutamate turnover may be associated with plasticity.

Keyword Control

Also included are genes as positive controls. These genes are known to be expressed in the functioning CNS and may act as a diagnostic test to ensure that a hybridisation has worked.

Additional control genes are included on the array are so-called housekeeping genes. These genes are constitutively expressed in all cell types and their expression is expected to be constant in control and experimental RNA populations (Suzuki et al., 2000). Four housekeeping genes were included on the custom microarray; *glyceraldehyde 3-phosphate dehydrogenase* (*Gapdh*), *β-Actin*, *Peptidylprolyl isomerase A* or *Cyclophilin A* (*Ppia*) and *Hypoxanthine Guanine Phosphoribosyl Transferase 1* (*Hprt*) (see also section 3.7.6).

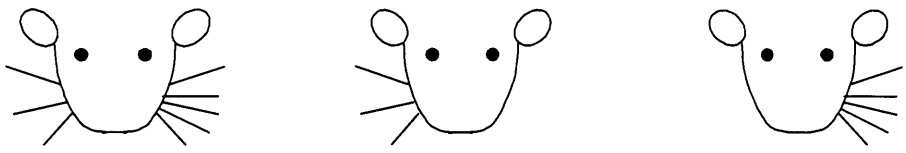
4.2.2 EXPERIMENTAL STRATEGY

The basic design of these experiments is a within-subject comparison in which different treatments are carried out on the two sets of mystacial vibrissae located on either side of the muzzle of a mouse. The contralateral barrel fields are removed, total RNA extracted and each sample labelled with a different fluorescent label. Both probes are then combined and hybridised to a custom microarray and genes differentially expressed

between the samples are identified, on the basis of a difference in the fluorescence generated at each spot.

As the experiment involves a within-subject comparison this limits the number of treatments that can be applied to each experimental unit (mouse) to two. However, the nature of whisker deprivation merits the application of three treatments ie if the gene expression profiles of chessboard deprived barrel cortices and undeprived barrel cortices are compared, differentially expressed genes may be associated with plasticity or may be differentially expressed purely as a result of changes in activity levels within the barrel cortex as a result of the whisker deprivation. Therefore, a third treatment, in which all whiskers are deprived (a treatment that does not induce plasticity or CRE-mediated gene expression) was included in the experiment design. This (all vibrissae deprived) treatment creates a problem in that rodents have only two sets of vibrissae, therefore a second animal was required in which one set of vibrissae are all deprived and the other are chessboard deprived.

As an additional level of control a third experimental unit is included in the design, where one set of vibrissae are undeprived and the other set are all deprived. This design concept is shown in figure 4.1.



(a) Chessboard vs Undeprived (b) Chessboard vs All deprived (c) All deprived vs Undeprived

Figure 4.1 The permutations of whisker deprivations for identifying differentially regulated genes following chessboard deprivation. Chessboard deprivation is represented by 3 whiskers, Undeprived (control) deprivation represented by 5 whiskers and an absence of whiskers indicates all whiskers deprived.
(a) One set of whiskers (always the animal's right hand side) is deprived in a chessboard pattern. The left hand side whiskers are undeprived.
(b) The right hand side whiskers are chessboard deprived while the left hand side whiskers are all deprived.
(c) The right hand side whiskers are all deprived and the left hand side whiskers are undeprived.

Another problem was the limiting amount of total RNA (average 2.4μg, section 3.2.2) available from the sampled cortical region containing the barrel cortex. Fluorescent probe would be prepared using the CyScribe™ direct labelling kit from Amersham Bioscience, UK (section 2.12.1.3) which recommends a minimum of 10μg of total RNA. Therefore, for a single hybridisation it was necessary to pool tissue samples from 4 animals prior to RNA extraction. The barrel fields that were to be compared were separately pooled and labelled (see figure 4.2). This method of pooling the total RNA, prior to labelling, meant that a single experiment which allows all distinct treatments to be compared against the other necessitated the use of 12 animals. Experiments were designed to compare the effects of deprivation at 1 day, 2 days, 4 days and 8 days (figure 4.3).

The experiment was to be replicated in independent experimental animals, in a form of replication peculiar to the use of two-colour cDNA microarrays – the 'Dye Swap' (Churchill, 2002, Kerr and Churchill, 2001). This replication involves reversing the fluorescent dyes used to label the RNA samples. If the first set of experiments are considered as XA and the replicated experiments as XB, then in XA all probes made from the pooled total RNA, derived from the animal's left hand barrel field were labelled with Cy3, whilst all right hand probes were labelled with Cy5 (section 2.3). In the replicated experiments, XB, this arrangement is switched so that left hand probes are labelled with Cy5 and right hand probes are labelled with Cy3. The dye-swap is shown schematically in figure 4.3, for a single time point, and is necessary to account for biases that may occur during labelling reactions or hybridisations where, for example a particular gene may preferentially label with Cy3. With a dye-swap replication this bias can be detected as the product of the replicated ratios will show a departure from the expected value of 1.

Consider two samples, x and y, that are compared in a dye swap. The replicated ratios of the measurement for gene$_a$ are designated R1 and R2

$$\text{Then } R1 = \frac{x_a}{y_a} = \frac{Cy3_{Xa}}{Cy5_{Ya}} \quad \text{and} \quad R2 = \frac{y_a}{x_a} = \frac{Cy3_{Ya}}{Cy5_{Xa}}$$

as the samples are identical

$$R1 * R2 = \frac{Cy3_{Xa}}{Cy5_{Ya}} * \frac{Cy3_{Ya}}{Cy5_{Xa}} = 1$$

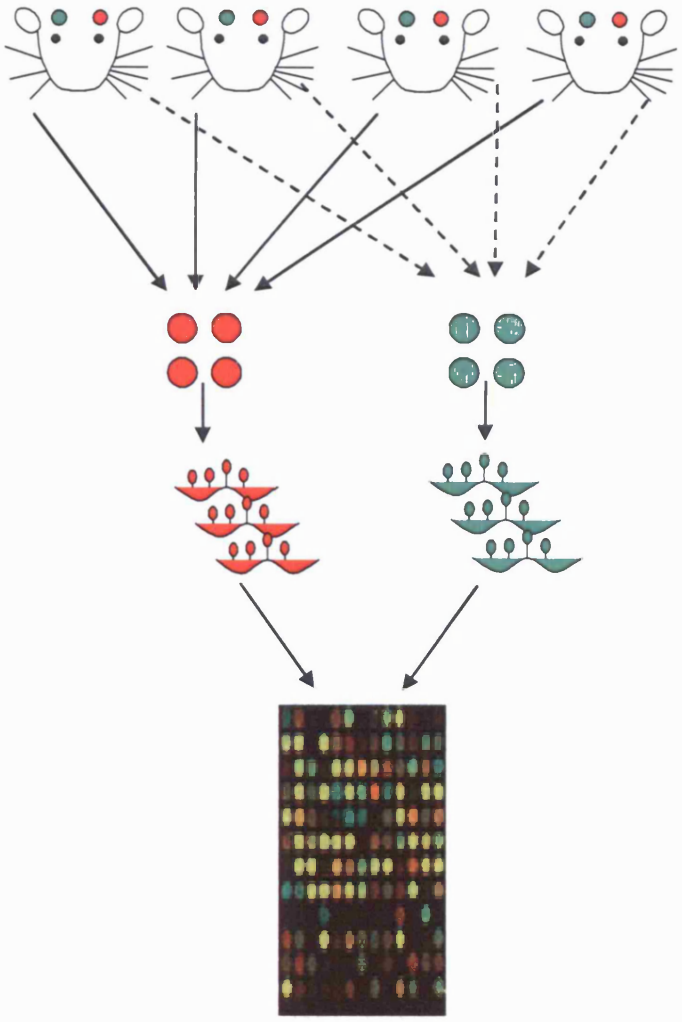


Figure 4.2 A schematic representation of the experimental design used to compare the gene expression profiles of chessboard deprived barrel cortices and undeprived barrel cortices. The whiskers of four animals are chessboard deprived (animal's right hand side, represented by 3 whiskers) and undeprived (animal's left hand side, represented by 5 whiskers). The contralateral barrel fields are removed and chessboard (red circles) and undeprived (green circles) barrel cortices are seperately pooled. Total RNA is extracted from each pool and is labelled with Cy3 and Cy5 fluorophores and hybridised to a custom microarray.

In reality, variation in the experiment would mean that the product of the replicated ratios would not be 1, but would be normally distributed and inconsistent genes would show a significant deviation from the mean. Values a certain number of standard deviations from the mean can then be removed as outliers (Quackenbush, 2002).

Each experiment consists of a single hybridisation using the pooled barrel cortices of 4 animals. The experiment code describes the deprivation applied to the vibrissae of the pooled animals;

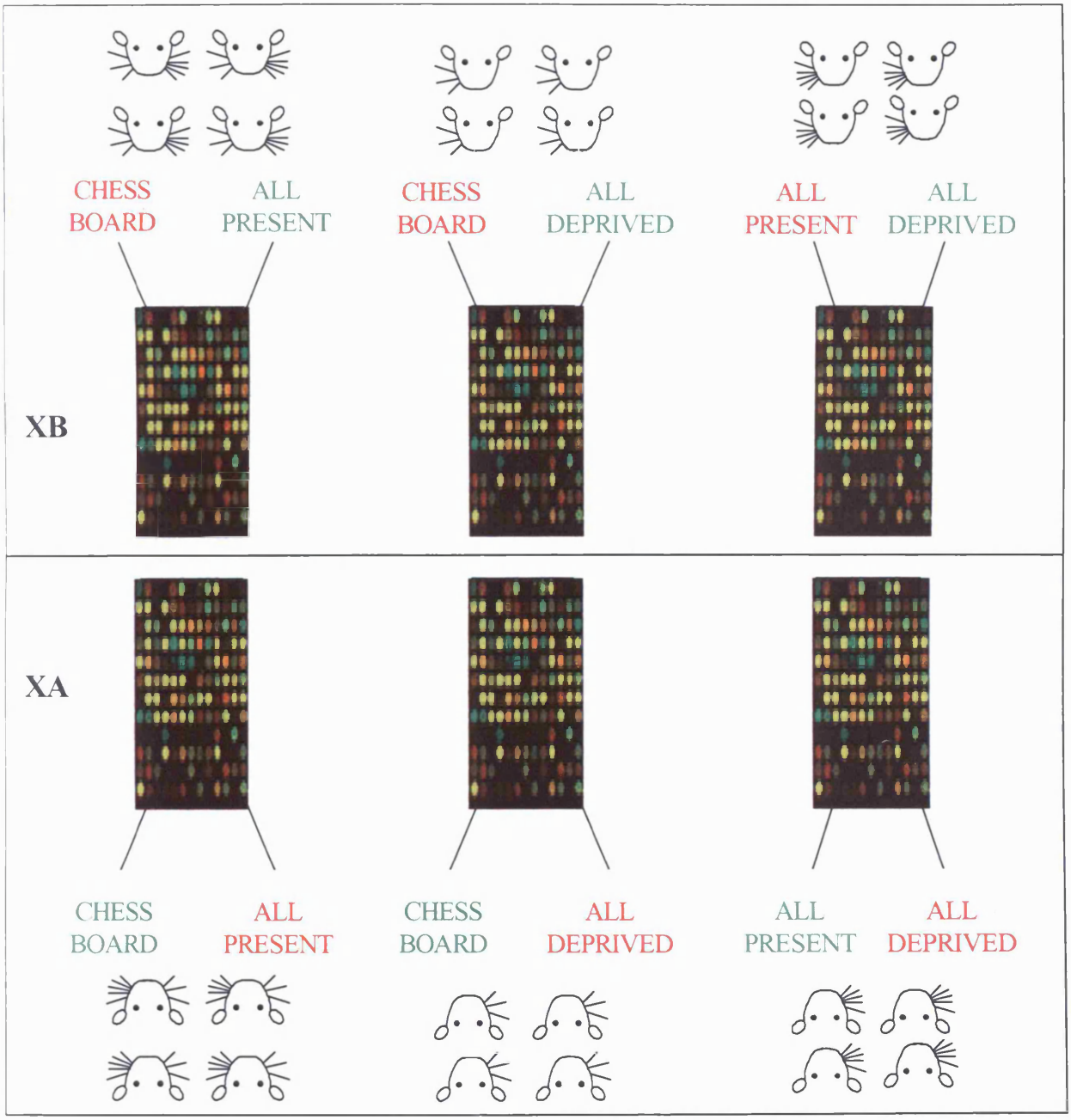


Figure 4.3 A schematic representation of the pairing of samples in a dye swap replication. For a single time point six hybridisations allow for the direct comparison of the gene expression profiles of barrel cortices subjected to three treatments with replication in independent biological samples. In the lower panel (XA) all the animals left hand barrel cortices are labelled with Cy3 dye and the left with Cy5. In the replicated experiment, XB, the labelling is reversed so that left is labelled Cy5 and right Cy3.

CB – Right Hand Side Vibrissae Chessboard Deprived, Left Hand Side Vibrissae Undeprived

DCB – Right Hand Side Vibrissae Chessboard Deprived, Left Hand Side Vibrissae All Deprived

DEP – Right Hand Side Vibrissae All Deprived, Left Hand Side Vibrissae Undeprived

XA represents the first set of hybridisations, XB hybridisations are the biological replications using a dye swap. The integer describes the length of deprivation either 1, 2, 4 or 8 days. Alternatively a symbol is given which pictorially describes the experiment (figure 4.1).

4.2.3 SAMPLE PREPARATION

The genetic background of the animals used in this series of experiments is described in section 2.2. At the time of deprivation animals were between 6-8 weeks old an age at which the anatomical development of the barrel cortex would have been complete and unaffected by whisker deprivation. Whisker deprivation procedures were carried out as described (section 2.3).

Following whisker deprivation whole brains were removed and barrel cortices were extracted as described (section 2.6). The brains were laterally bisected and sectioned (section 2.7.1-2.7.2). Sections were stained using the cytochrome oxidase reaction in order to visualise the posteromedial barrel field (PMBSF) in layer IV of the cortex (section 2.7.3). Typical sections are shown in figure 4.4 and reveal that in all cases the PMBSFs were successfully removed. The hole left by the borer used in the extraction is in the position usually occupied by the PMBSF (compare figures 4.4 i and 4.4 a- h) and although some barrels are visible these represent the short whiskers on the snout unaffected by experimental procedures. Using this quality control measure ensured that microarray data was not affected by sampling error.

The production of high-quality probe was dependent on the quality of the total RNA extracted from cortical tissue. Extraction of total RNA was carried out as described (section 2.8.2).

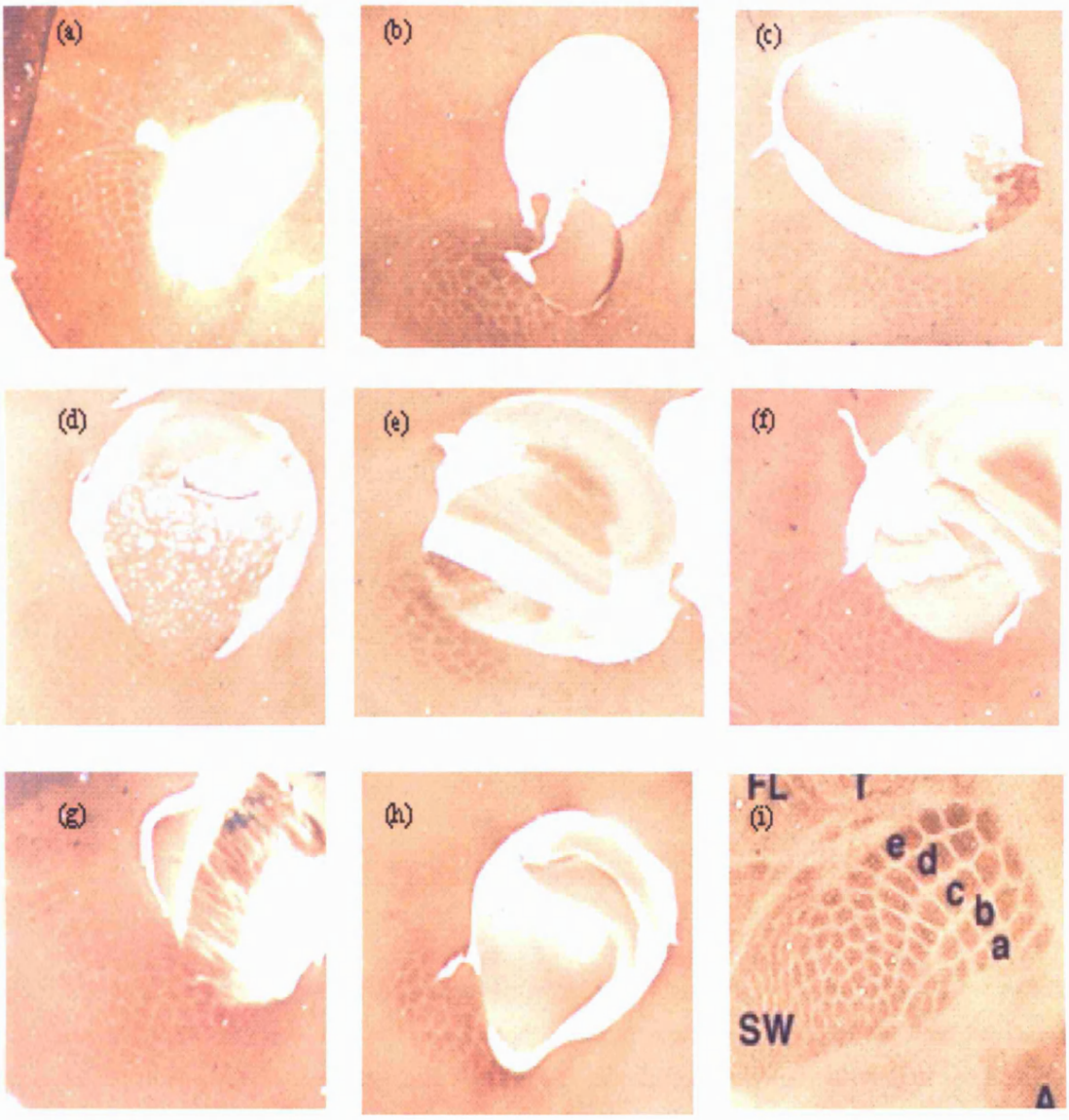


Figure 4.4 Examples of cytochrome oxidase stained sections of layer IV of the mouse somatosensory cortex. Figure (i) shows a complete barrel field. The barrels representing the large mystacial vibrissae are arranged in rows labelled a-e. These larger barrels form the posteromedial barrel subfield (PMBSF). SW refers to smaller barrels representing the short whiskers on the animals snout. In figures (a) – (h) the PMBSF has been successfully removed prior to microarray analysis.

All total RNA samples extracted from the pooled tissue removed from the cortex of mice were subjected to quality control measures prior to microarray probe preparation (section 3.2.2). Every sample appeared undegraded when visualised on agarose gels (figure 4.5), with the 28S and 18S ribosomal bands appearing well defined without the smearing indicative of degraded samples (see figure 3.3 for an example of degraded total RNA).

The OD260 : OD280 measurements ranged between 1.7 and 2.1 for the total RNA samples indicating pure samples free from contaminants such as proteins or phenol (section 3.2.2).

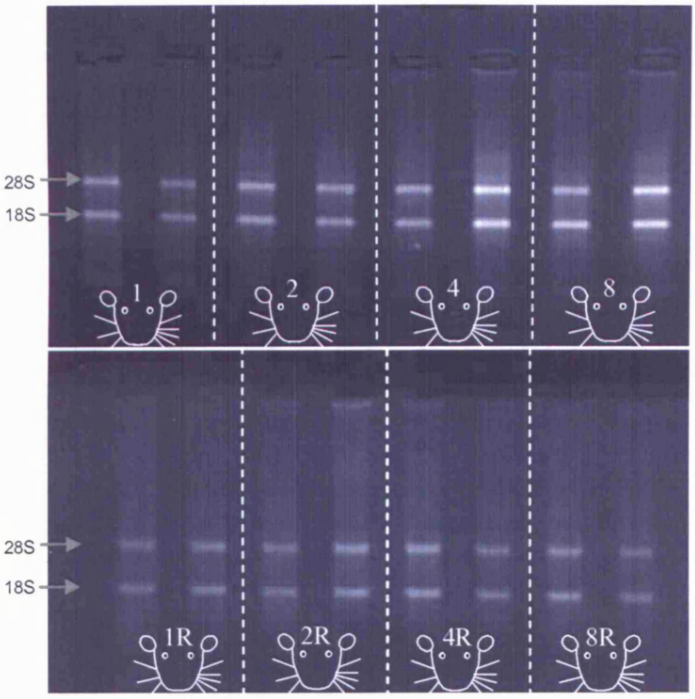


Figure 4.5 Agarose gels of total RNA samples extracted from tissue removed from the cortex of mice. Bright, well defined 18S and 28S ribosomal bands indicate intact (undegraded) samples suitable for fluorescent labelling prior to microarray analysis. For a description of experimental symbols see figure 4.1. Gridded lines separate samples paired in hybridisations. Within each grid, samples to the left were obtained from animal's left side cortex and those to the right were from right side cortices.

Total RNA samples (10μg) were fluorescently labelled using the protocol described (section 2.10.3). The labelled probe generated from every total RNA sample was analysed using agarose gel electrophoresis (section 2.10.3). Typical examples of the gel images are shown in figure 4.6. All samples formed intense fluorescent smears suggesting efficient probe generation. Equal incorporation of both the Cy3 and Cy5 fluorophores had occurred as can be seen from the relative intensity in the two channels. No obvious differences in the size range of the probes, generated using Cy3 or Cy5, was apparent.

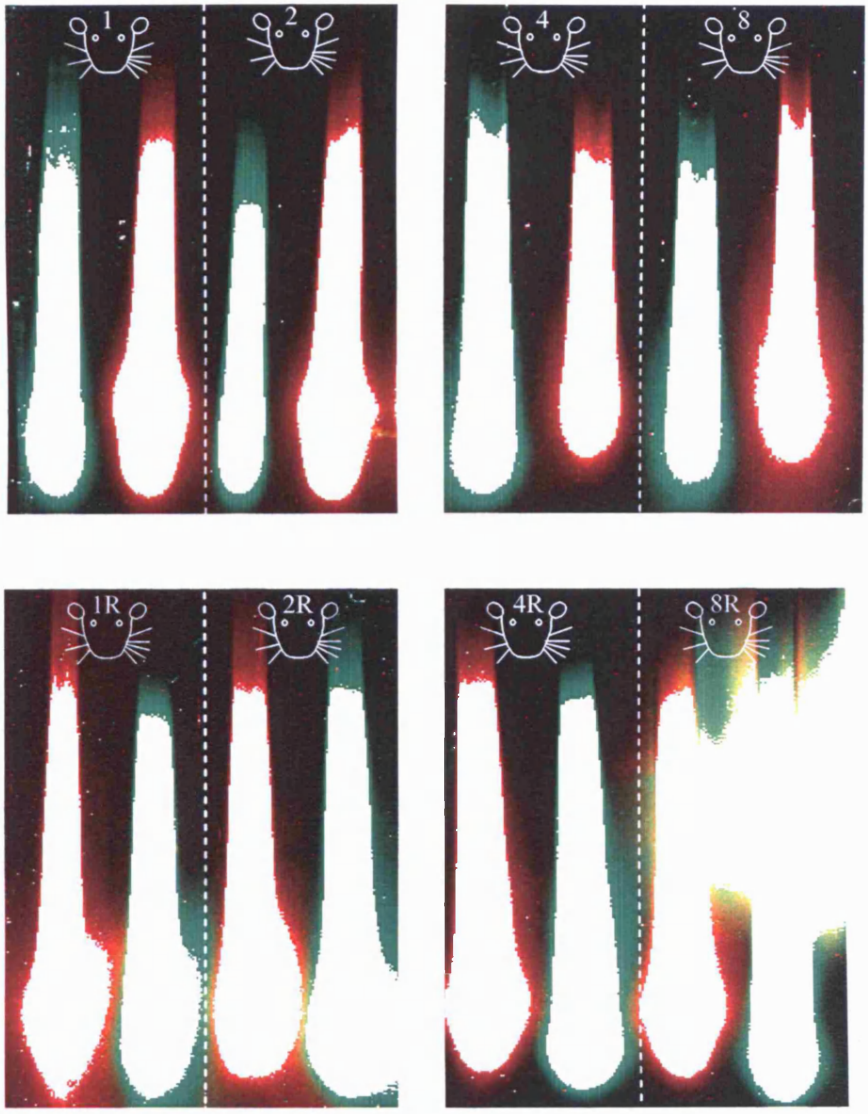


Figure 4.6 Examples of agarose gels showing successful generation of Cy3 (green) and Cy5 (red) labelled probes for microarray hybridisation. The gridded lines separate pairs of probes that will be combined in hybridisations. The left hand probes in each pair were generated from pooled total RNA (see figure 4.3) extracted from left hand cortices while the right hand probes were from right hand cortices. The symbols represent the form and length of time of vibrissae deprivation (figure 4.1). Notice that all right hand cortices are labelled with Cy3 and left hand side cortices are labelled with Cy5 in the upper panel. In the lower panel this labelling is reversed which is the basis of a 'Dye Swap' experiment.

Following quality control labelled probes were combined and hybridised (section 2.16.3.1) to custom microarray slides (section 2.15.1) which were scanned as described (section 2. 17).

4.2.4 DATA ANALYSIS

Following hybridisation and scanning each microarray image was analysed and data extracted. Figure 4.7 provides the reader with a reminder of the process of data extraction that ends with identification of differentially expressed genes.

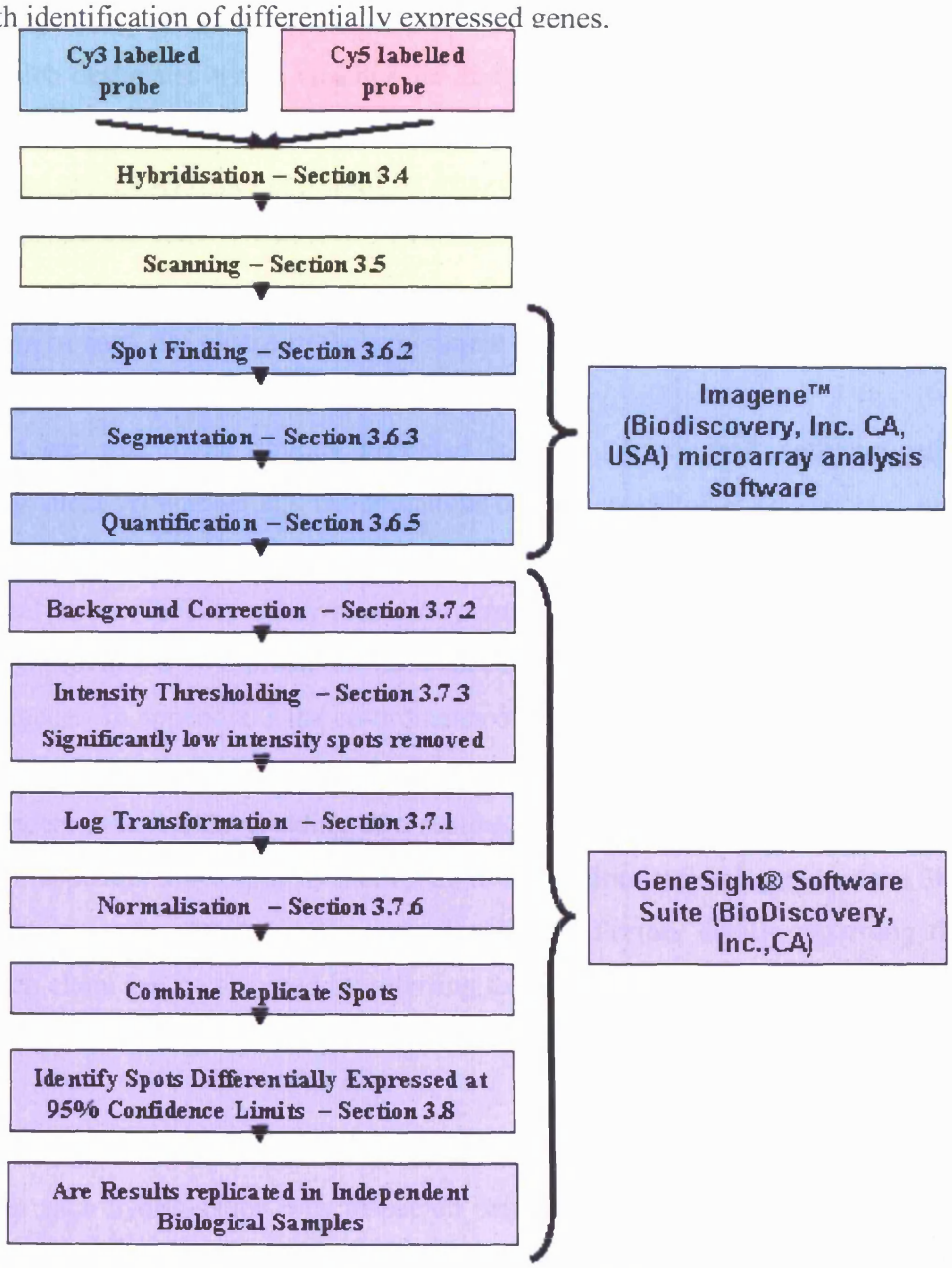


Figure 4.7 A flowchart describing the sequence of events that are involved in identifying differentially expressed genes from a scanned microarray image. Most steps contain a reference to the relevant section where the procedure is optimised and / or described further. Also included are details of the software used to carry out each step.

Each procedure is accompanied by a reference to a section of this thesis in which the procedure was optimised and / or described further. Suffice to say that every microarray was analysed in exactly the same

Raw data from all hybridisations is provided in appendix 2. The data is available in two forms;

a) for each hybridisation two TIFF files are provided, one for each channel of the image. Each TIFF file is assigned a name which describes the experiment it relates to followed by a letter which designates which channel the image represents. Importantly 'A' files represent the Cy3 channel and 'B' files the Cy5 channel.

b) also provided are the text files produced by Imagene™ (Biodiscovery, Inc. CA, USA) microarray analysis software. These text files contain the quantified fluorescence of each spot in addition to the co-ordinates of each spot, the spot ID and flagging information. The description of each file relates to the experiment name and the 'A' / 'B' information is as described for the TIFF file.

The text files are able to be directly imported into the GeneSight® software suite (BioDiscovery, Inc.,CA) without any manipulations but can conceivably be imported into any microarray analysis software that supports the .txt file format.

In order to analyse the TIFF files any microarray image analysis software will require a template which provides a co-ordinate for each spot in addition to information relating to the ID of the gene. In appendix 3 the co-ordinates of each spot are provided in the first four columns headed Meta Row, Meta Column, Row and Column. In order to visualise how these headers describe the gridding of a custom microarray refer to figure 3.15. In column five of appendix 3 is a spot ID that refers to the position of each gene on the 384 well plate containing the cDNA used to print the array. Further details regarding the identity of each clone can be obtained by referring to appendix 1, column 8 in which the spot ID is tied into all the required clone information.

4.3 RESULTS

The data from each hybridisation was inspected using RI plots (figure 4.8). For each hybridisation the raw and transformed RI plots are provided, in order to judge the effectiveness of data transformation. In all cases the log2 (Cy3/Cy5), or R, is plotted as a function of the log10 (Cy3*Cy5), or I, for every spot on the microarray.

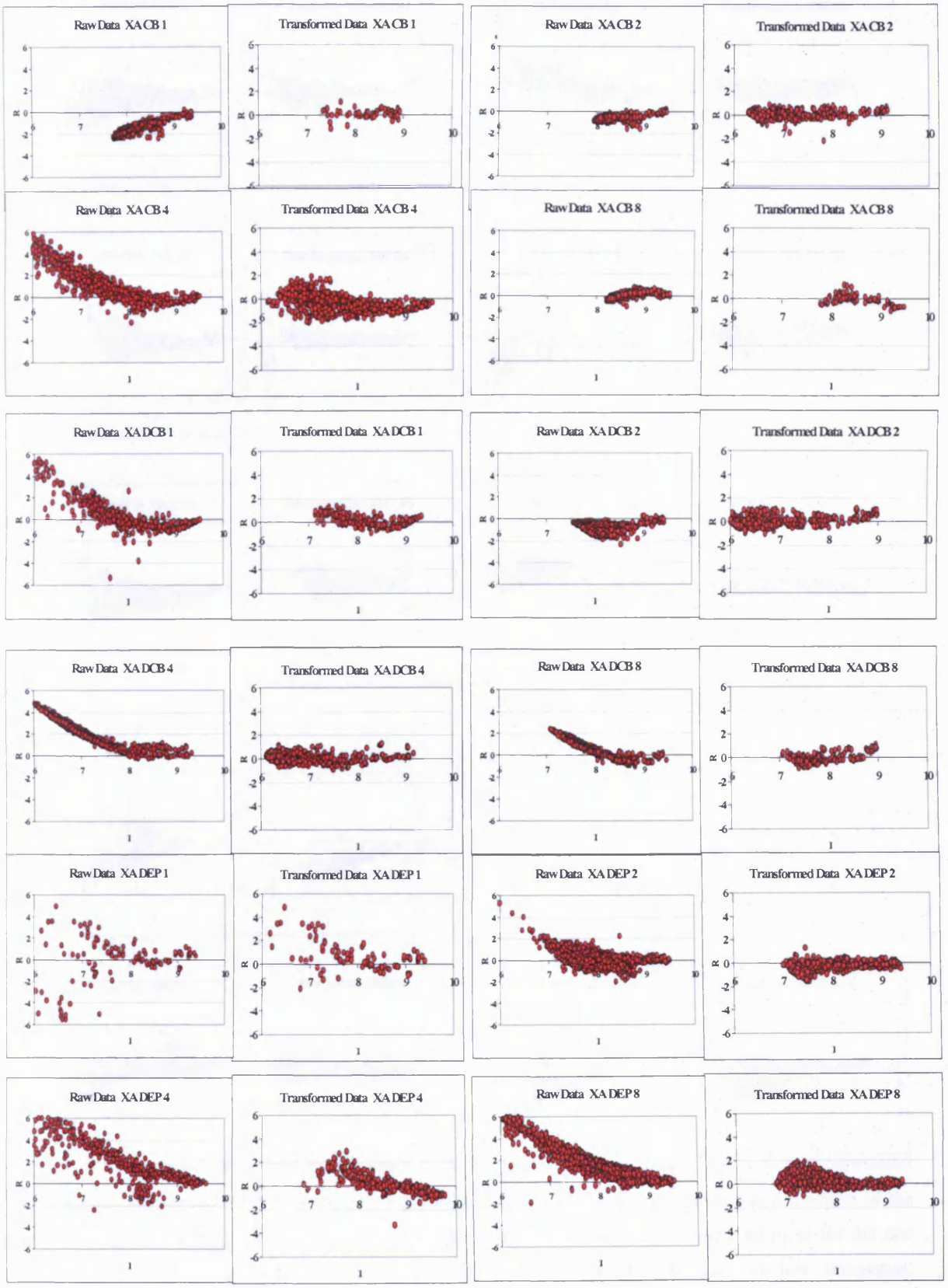


Figure 4.8 Raw and transformed data RI plots for experiments XA. For experimental descriptions see section 4.2.2 (continued overleaf)

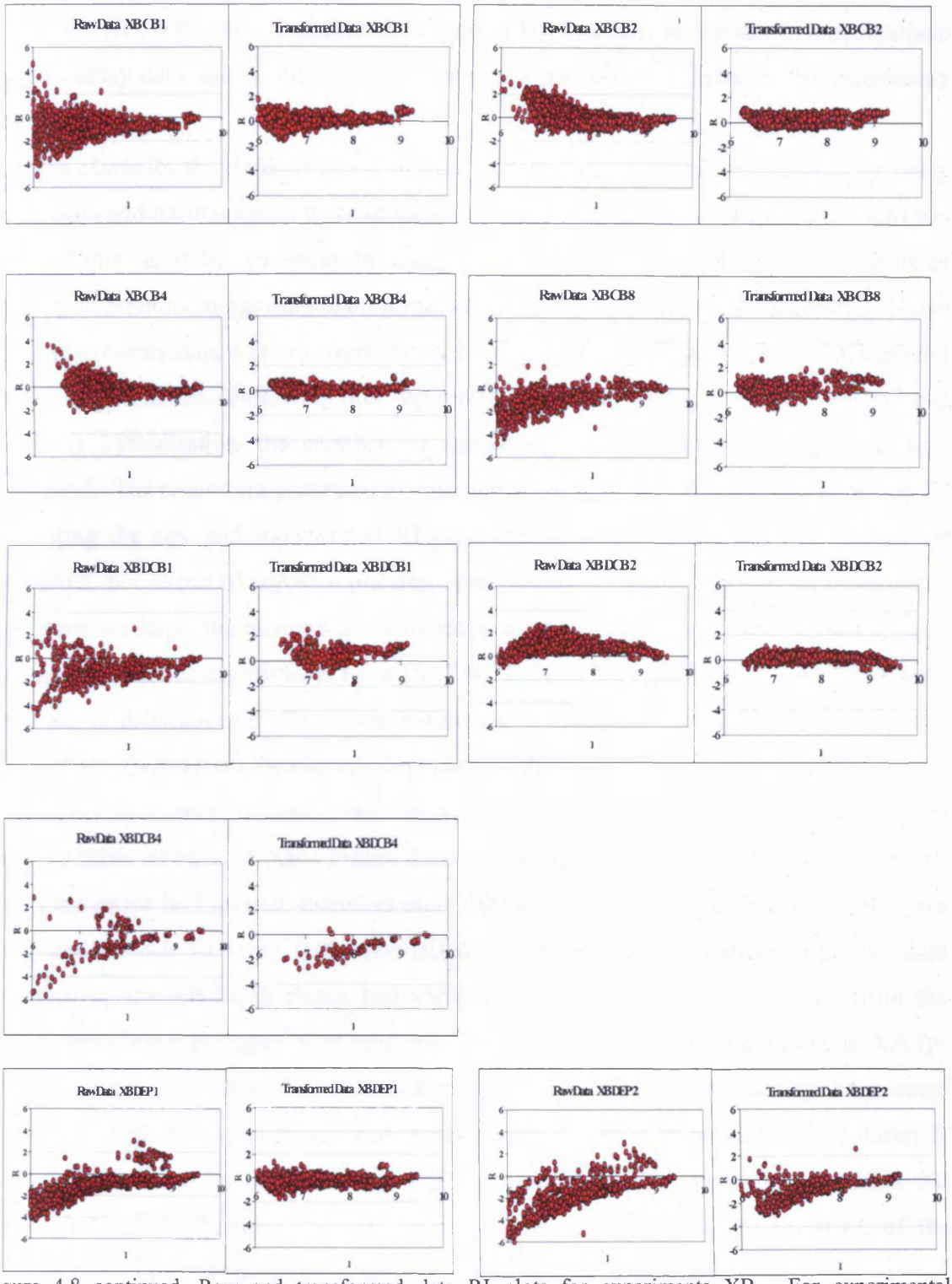


Figure 4.8 continued. Raw and transformed data RI plots for experiments XB. For experimental descriptions see section 4.2.2. The RI plot represents the $\log_2$ (ratio), y- axis, plotted as a function of the $\log_{10}$ (Cy3 intensity * Cy5 intensity), x-axis, for each spot on the array. Note many RI plots for the raw data show a distinct curvature indicating a systematic variation in the ratio at low intensities. Transformation of the data attempts to remove this curve, successful in some cases, eg XBDEP1, not in others XBDEP2.

Prior to transformation, the raw data RI plots for many experiments, eg XACB4, XADCB1, XBDEP1 show a distinct curvature at low intensity levels that is characteristic of microarray data and is indicative of systematic variations inherent in the microarray process.

In these examples the transformation of data ie background correction, removal of noisy data points and normalisation by housekeeping genes present on the array (figure 4.8) has removed this variation. However, in many cases eg XBDEP2, XBDCB4 the curvature of the RI plot remains, suggesting that normalisation has not been entirely successful. Noise in the microarray data was removed by only using spots significantly above background levels (section 3.7.3). Following this step replicate spots of each gene were averaged and for each hybridisation the number of genes significantly above background was calculated. The results are presented in column two of table 4.1, and can be visualised by comparing the raw and transformed RI plots (figure 4.8) which show the number of individual, not averaged across replicates, spots removed by this process. It is apparent that there are large fluctuations in the number of spots significantly above background amongst the various experiments from 15 (XA CB 1) to 327 (XB DCB 2). As there were no apparent differences in the probes hybridised to each microarray (section 4.2.3), the idea that the microarrays themselves were causing this variation was investigated.

Microarray slides were printed in batches with experiments labelled XA using a different batch to those designated XB. Figure 4.9a reveals that significantly greater numbers of genes are above background, therefore more data was collected, in the batch of XB slides compared to XA (p-value 0.03, two-tailed Student's t-test). Furthermore the data collected on the XB batch shows less variation than the XA, as can be seen from the significantly lower average CV of replicated spot ratios for XB batch compared to XA (p-value < 0.01, two-tailed Student's t-test, figure 4.9 b). For each hybridisation the average CV of the replicated, significantly above background, spots is presented in column 5, table 4.1. It could be possible that, with practice (XA experiments were carried out before XB), the techniques involved in performing hybridisations eg lowering of the cover slip (section 3.4) were being performed with greater efficiency.

EXPERIMENT	NUMBER OF GENES SIGNIFICANTLY ABOVE BACKGROUND	AVERAGE RATIO OF REPLICATED SPOTS- RAW DATA	AVERAGE RATIO OF REPLICATED SPOTS- TRANSFORMED DATA	AV. CV OF REPLICATED SPOT RATIOS	NUMBER OF GENES DIFFERENTIALLY EXPRESSED
XA CB 1	15	-1.64	0.08	-1.26	2
XA CB 2	116	-0.55	0.10	1.93	2
XA CB 4	202	4.05	-0.29	-1.35	31
XA CB 8	25	-0.01	0.04	-1.17	4
XA DCB 1	56	3.76	0.24	7.28	3
XA DCB 2	146	-0.76	0.04	0.14	14
XA DCB 4	161	2.02	0.10	6.05	10
XA DCB 8	45	0.69	-0.004	-0.14	4
XA DEP 1	29	0.84	0.80	0.16	11
XA DEP 2	145	0.05	-0.26	-0.67	8
XA DEP 4	97	5.59	0.21	0.30	32
XA DEP 8	257	3.26	0.40	-0.27	4
XB CB 1	267	1.21	0.17	0.34	26
XB CB 2	265	0.01	0.35	-0.18	17
XB CB 4	217	0.33	0.13	-2.34	3
XB CB 8	138	4.56	0.34	-0.14	37
XB DCB 1	105	2.26	0.56	0.09	15
XB DCB 2	327	-0.78	0.25	0.32	5
XB DCB 4	23	2.78	-1.18	-0.28	3
XB DCB 8	Experiment Fail				
XB DEP 1	270	2.52	-0.22	-0.22	8
XB DEP 2	160	5.47	-0.56	-0.94	28
XB DEP 4	Experiment Fail				
XB DEP 8	Experiment Fail				

Table 4.1 A summary of the results from hybridisations to a custom microarray. This data was compiled after averaging the Cy3/Cy5 ratios of the four replicates of each gene on each microarray. Identification of genes significantly above background was carried out as described in section 3.7.3, the average CV (std dev / mean) of replicated is provided as a measure of the variation amongst replicates. Differentially expressed genes were identified by noise sampling as described (section 3.8).

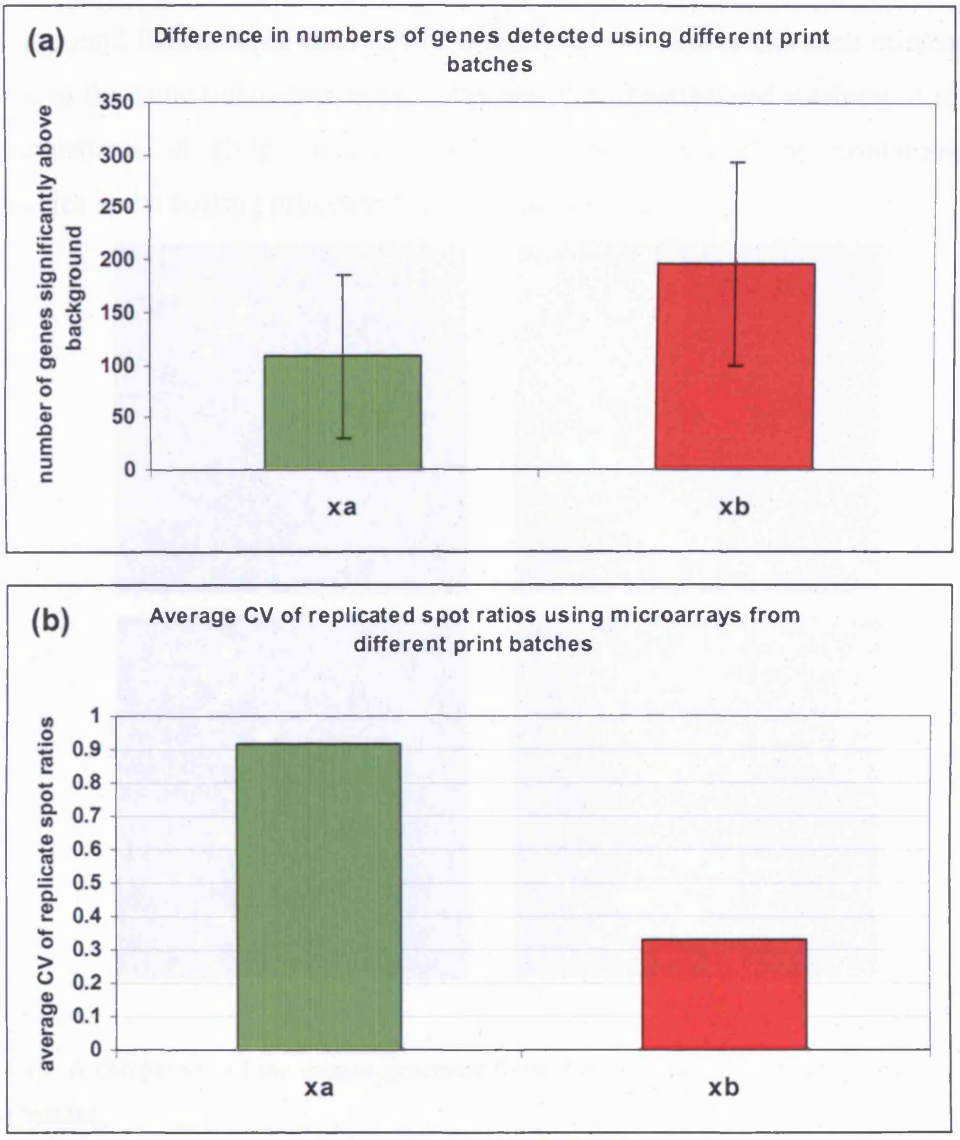


Figure 4.9 a and b. A comparison of the data generated using slides from different print batches.

a) The average number of genes significantly above background from hybridisations carried out in experiments XA compared with XB. The experiments used slides from different print runs and different manufacturing lots and show significant differences in the amount of data generated ($p=0.03$, 2-tailed Student's T-Test).

b) A comparison of the quality of data generated using slides from different batches. The CV is calculated for each gene using the mean / standard deviation of the replicate spots on the array, which is then averaged to provide a CV for each hybridisation (see table 4.1). These CV values are averaged across each print batch ie XA or XB to provide a measure of the inherent variation when using slides from different print runs. The lower CV for the XB experiments suggest significantly lower variability in the data generated using this batch of microarrays ($p=0.01$, 2-tailed Student's T-Test).

It is equally probable that the XA batch slides were of a lower quality than the XB. Indeed, the latter appears to be the case as the majority of slides used in XB showed far

less background fluorescence than XA (see figure 4.10). Given that each microarray was subjected to the same treatments eg printing, pre-hybridisation and washing, it is possible that deficiencies in slide quality could have been caused by contamination or inadequacies in the coating procedure during manufacture.

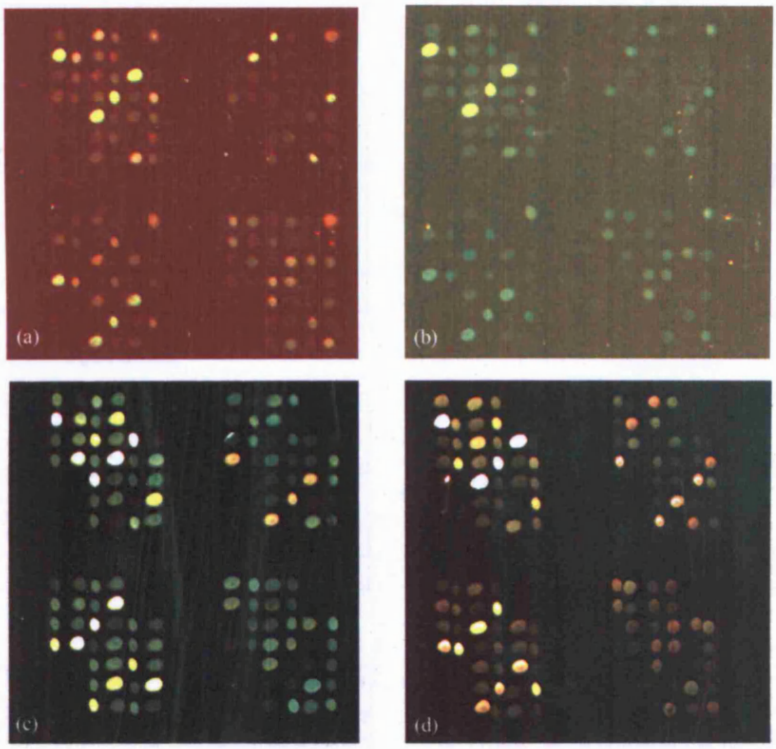


Figure 4.10 A comparison of the images generated from hybridisations carried out on microarrays from separate batches.

a) and b) are from experiments XA CB 2 and XA DEP 2 respectively.

c) and d) are from experiments XB CB 8 and XB DEP 2 respectively.

Note the lack of background fluorescence in c) and d) indicative of better quality images. Notice how the lower background accentuates the intensity of the spots.

The process of identifying significantly differentially expressed genes and the identification of those significant genes common amongst replicate experiments was carried out using the Confidence Analysis tool within GeneSight® software suite (BioDiscovery, Inc.,CA).(section 3.8).

In order to reduce the number of false positives generated, for a gene to be classed as differentially expressed between two treatments it was required to be significant in both replicates of the same experiment ie the gene must be differentially expressed in both XA and XB.

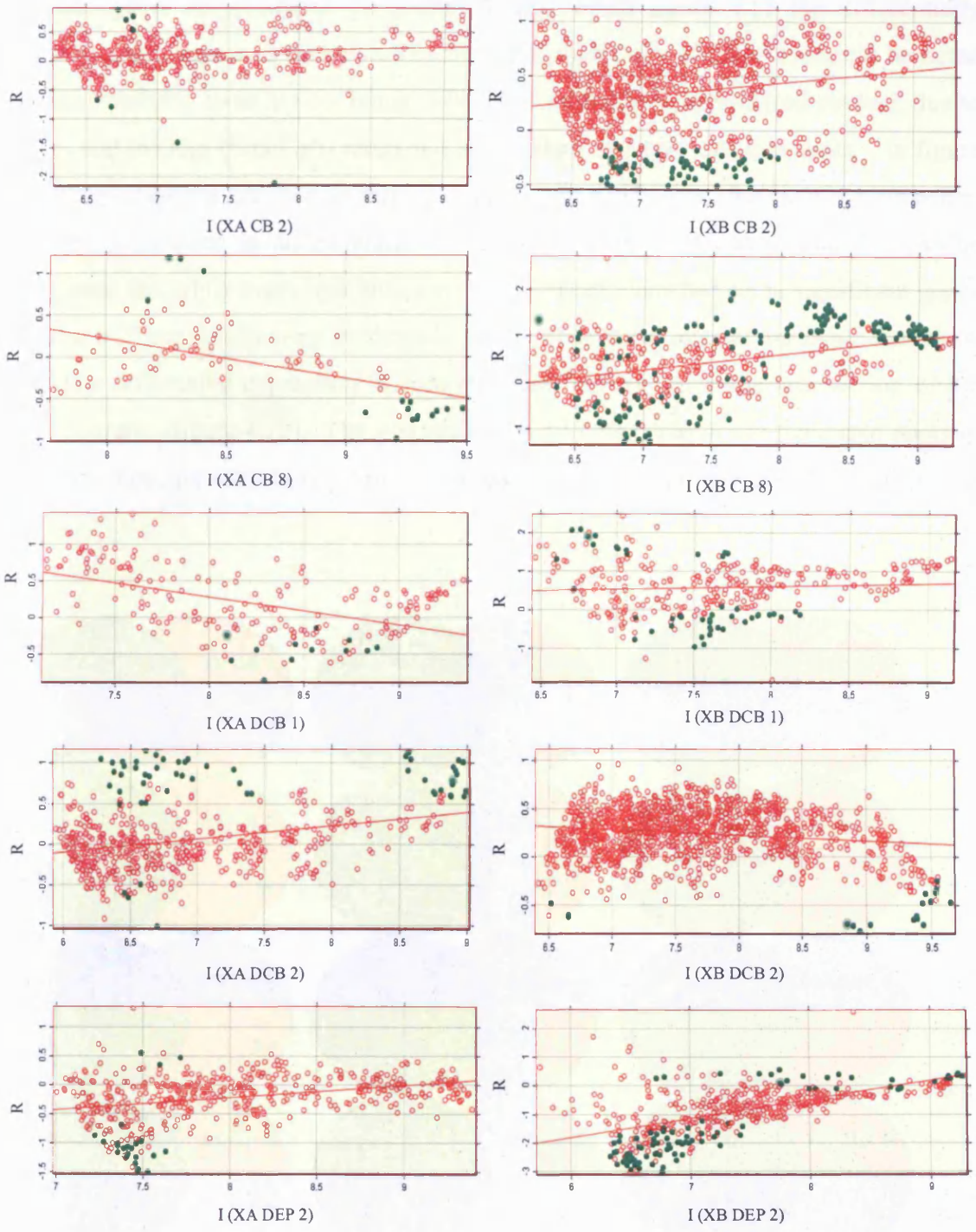


Figure 4.11 RI plots for replicate hybridisations that have differentially expressed genes in common. Note that data from every spot significantly above background is plotted ie replicate spots are not combined. Green data points indicate genes significantly differentially expressed. The differentially expressed genes appear as outliers from the majority of data points. Not all outliers are being called differentially expressed as the replicate spots fail to generate similar ratios.

The RI plots for every pair of replicate experiments that have differentially expressed genes in common are presented in figure 4.11, in which significantly differentially

expressed genes are indicated by green circles. From figure 4.11 the differentially expressed genes appear as outliers away from the main body of data. It should be noted that many outlying spots are not being called differentially expressed (red circles), due to the fact that the $\log_2$ (Cy3/Cy5) ratios fail to correlate amongst replicate spots. In figure 4.12, Venn diagrams are used to visualise the numbers of genes, for each hybridisation, that were determined as differentially expressed. From the Venn diagrams it can be clearly seen that while every hybridisation (that worked) identified some significant genes very few of these results were replicated. Genes of greatest interest are those that show significant differential expression in both replicates ie they are in the intersection of the Venn diagrams (figure 4.12). The average $\log_2$ (Cy3/Cy5) ratio of the replicated spots in the intersection, for each biologically replicated experiment, are presented in figure 4.13 a-e.

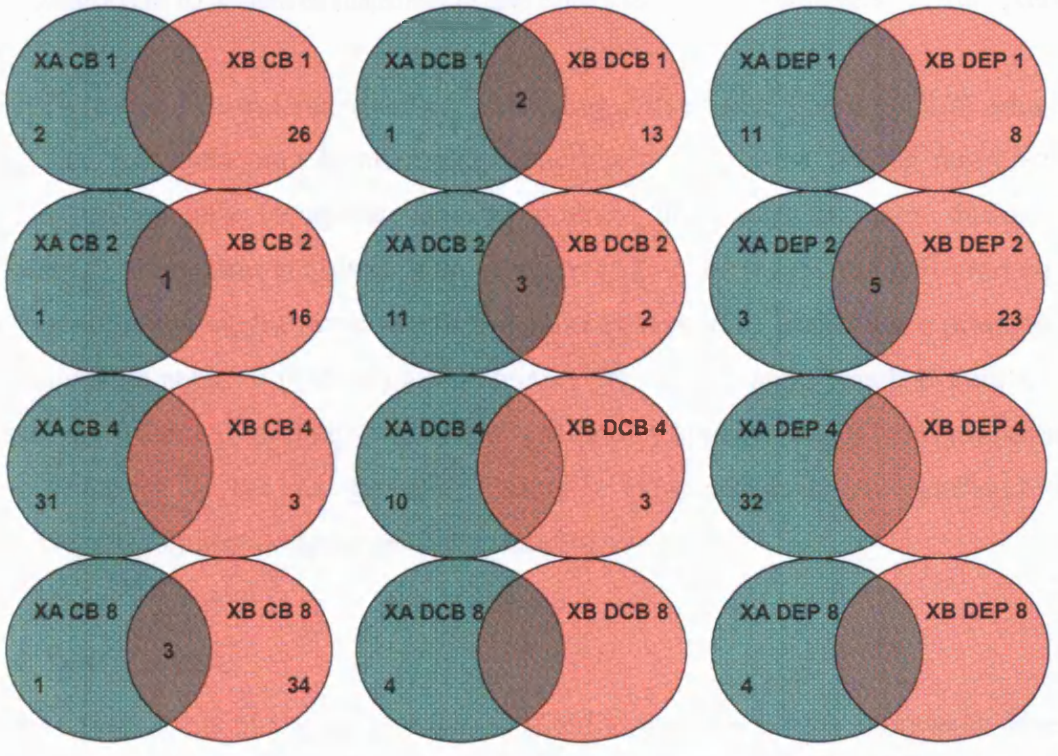


Figure 4.12 A Venn diagram illustrating the number of significantly differentially expressed genes identified for each microarray experiment. The intersection provides the number of genes that appear significant in an independent biological replicate of the hybridisation. The figures outside the intersection relate to the numbers of significant genes for that experiment that failed to be replicated. Failed experiments does not have any associated figures.

EXPERIMENT	GENE ID	NIA 15K CLONE SET ID	PUTATIVE GENE ID
XA/ XB CB 8	D17	H3054H04	KCNN4
XA/ XB CB 8	H19	H3139C10	Sstr1
XA/XB DCB 8	A4	H3008B08	purα
XA/XB DCB 8	E21	H3121H02	CALM2
XA/XB DCB 8	O19	H3127F12	Macf
XA/XB DEP 2	O15	H3009E03	activin receptor

Table 4.2 provides information relating to the identity of the genes that are significantly differentially expressed, in biologically replicated microarray experiments. Gene ID refers to the 384 well position of the respective clone, NIA ID refers to the specific clone and its identification within the NIA 15K clone set. The putative gene ID is based on similarities of each clone's sequence with sequences of known genes.

The replicated experiments involved performing a dye swap (section 4.2.2) which is a technical procedure used in microarray experiments to account for the fact that some genes preferentially incorporate one or other of the fluorescent dyes during probe labelling. Such genes are identifiable in figure 4.13 as the $\log_2$ (Cy3/Cy5) ratio will, for both replicates, be in the same direction ie either both positive (indicating a preference for Cy3) or both negative (indicating a preference for Cy5). Genes (marked with a star in figure 4.13) that are resistant to this effect will display opposite $\log_2$ (Cy3/Cy5) ratios.
The identities of the six genes differentially expressed that are significant under experimental replication are presented in table 4.2.

4.4 DISCUSSION

As the main focus of this project lies in identifying genes whose expression levels are regulated following chessboard deprivation, perhaps the most interesting results are those from XA/XB CB 8 and XA/XB DCB 2, where significantly differentially expressed genes, are observed when the experiment is replicated (figure 4.12), a number of which are resistant to preferential dye incorporation (figures 4.13 b and d respectively).

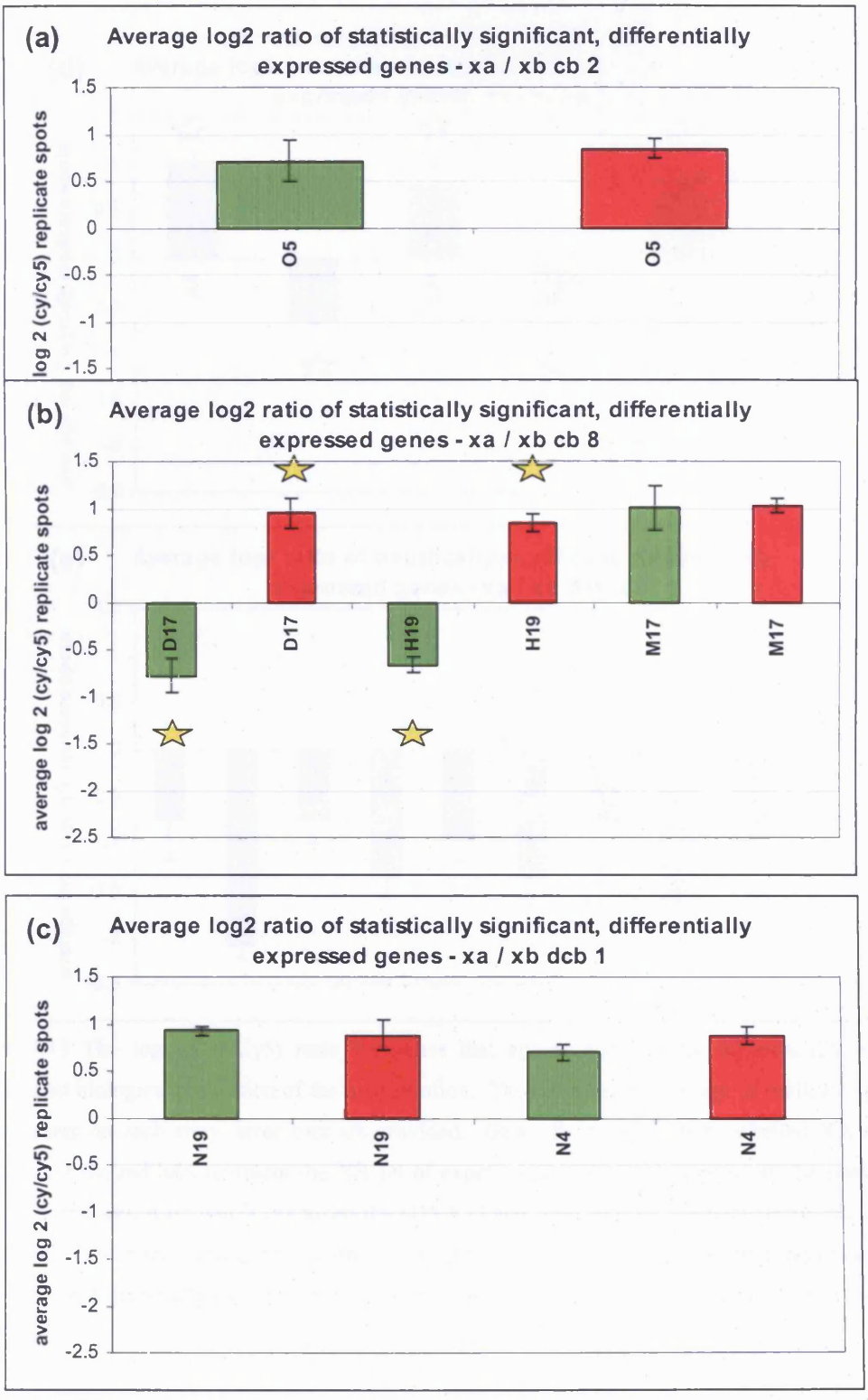


Figure 4.13 (continued overleaf)

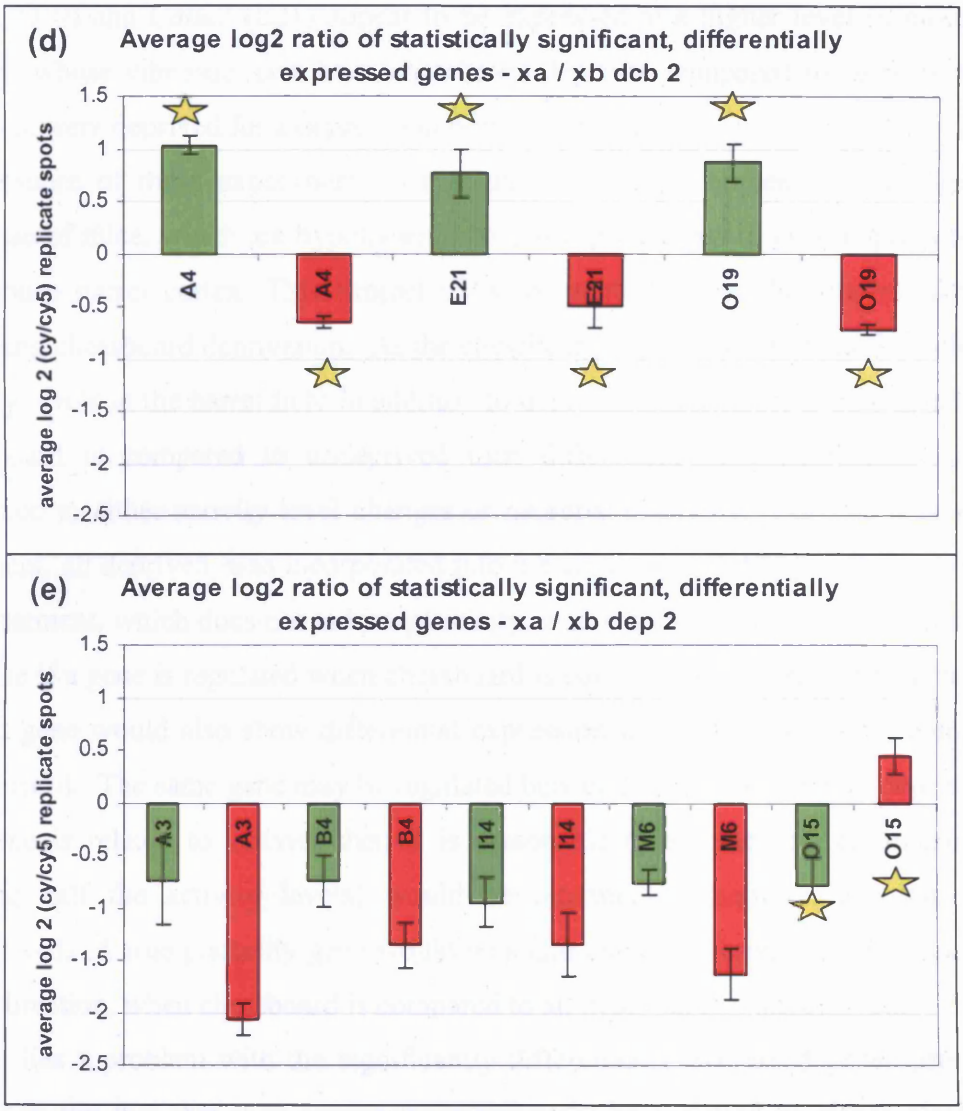


Figure 4.13 The $\log_2$ (Cy3/Cy5) ratio for genes that appear significantly differentially expressed in independent biological replication of the hybridisation. The ratios are an average of replicate spots (four in total) present on each array, error bars are provided. Genes from experiments labelled XA are coloured green, while the red bars represent the XB set of experiments. The code represents the position of each gene on a 384 well plate which contained the cDNA clones used to print the microarray (table 4.2). Bars marked with a star represent genes whose $\log_2$ (Cy3/Cy5) ratios oppose one another suggesting that these genes are not differentially expressed because of preferential incorporation of one or other of the fluorescent dyes.

As can be seen from figure 4.13 b the genes *Kcnn4* (D17) and *Sstr1* (H19) are down regulated in chessboard deprived barrel cortices compared to undeprived barrel fields when the treatment is applied for eight days.

From figure 4.15 d, three genes, *purα* (A4), levels *macf* (microtubule-actin cross-linking factor, O19) and *Calm2* (E21) appear to be expressed at a higher level in mouse barrel cortices whose vibrissae have been chessboard deprived compared to cortices where all vibrissae were deprived for a deprivation period of 8 days.

The essence of these experiments is that three separate treatments are applied to the vibrissae of mice, which are hypothesised to result in changes in gene expression within the mouse barrel cortex. This project aims to identify genes that are regulated only following chessboard deprivation. As the chessboard deprivation will result in changes in activity levels in the barrel field in addition to neuronal plasticity, it is conceivable that if chessboard is compared to undeprived then differential gene expression cannot be attributed to either activity level changes or neuronal plasticity. For this reason a third treatment, all deprived, was incorporated into the experimental design. It was hoped that this treatment, which does not induce plasticity, would control for changes due to activity levels ie if a gene is regulated when chessboard is compared to all present then an activity related gene would also show differential expression when all present was compared to all deprived. The same gene may be regulated between chessboard and all deprived but if the gene is related to activity then it is reasonable to assume that chessboard levels (having half the activity levels) would be intermediate between all deprived and undeprived. A true plasticity gene would be significantly differentially expressed, in the same direction, when chessboard is compared to all deprived and undeprived.

Herein lies a problem with the significantly differentially expressed genes presented in table 4.2; the fact that they cannot unambiguously be assigned as either plasticity, or activity, related genes as they do not appear as significant in other comparisons.

The reason for this is unclear but may be because;

a) the genes are false positives and not truly differentially expressed

b) the genes are false negatives in other comparisons

Problem a) may be attributed to the method used to identify statistically significant differentially expressed genes. Here a noise sampling method was used (section 3.8) that is proposed to overcome the problem of genes with very high or very low intensity levels appearing to be differentially expressed (Draghici et al., 2003). The method weights the ratios of all genes according to the ratios of other genes with similar intensities. While a number of differentially expressed spots are observed through a wide range of intensity levels, many are being called differentially regulated at the extreme tails where inherent noise has the greatest impact (figure 4.11). Increasing biological replication would allow

for a more thorough statistical analysis of microarray data which would reduce the number of false positives (Churchill, 2002).

Problem b) may not be completely resolved by replicating experiments. These experiments have been compromised by low quality arrays eg it has been shown that microarrays from different print batches can seriously alter the number of genes detected (figure 4.10). Furthermore, the number of genes detected above background shows a large variance even in slides from the same batch and complete failure of hybridisation is not uncommon (table 4.1). It seems therefore that low quality microarrays are an inherent nuisance variable that must be dealt with, otherwise replicated experiments will be subject to the same variability. The print batch problem can be solved by printing all the required slides for a series of experiments, mixing the batches and then randomly assigning each microarray to a particular hybridisation (Churchill, 2002). Another possible way of dealing with low quality slides is to dispense with the custom microarray and use slides with more spots. Here, the impact of low quality slides would be less than for the custom microarray as significant numbers of data points should still remain after filtering low intensity spots.

To summarise, this chapter identified differential gene expression in the mouse barrel cortex following vibrissal deprivation. However technical replication of experiments failed to confirm these findings. In the next series of experiments an RNA amplification technique which will increase biological replication and hence statistical power to detect differential gene expression will be used. Furthermore, these experiments have shown up the limitations of using a set of housekeeping genes for normalisation as a number of arrays still show evidence of systematic variation after transformation (figure 4.8). In the next series of experiments NIA 15K microarrays, constructed from 15,000 transcripts, will be used. These arrays are amenable to more thorough normalisation procedures (Quackenbush, 2002) and will provide better coverage of the mouse transcriptome

CHAPTER FIVE

HIGH DENSITY MICROARRAY ANALYSIS OF GENE EXPRESSION IN THE BARREL CORTEX FOLLOWING VIBRISSAL DEPRIVATION

5.1 AIM

This chapter describes the gene expression profiling, using microarray technology, of mouse barrel cortices whose corresponding vibrissae were subjected to three distinct treatments. The main aim is to identify genes, that are differentially expressed in response to a particular kind of vibrissal deprivation (chessboard), that induces neuronal plasticity within the barrel field. The effects of vibrissal deprivation, over a period of 8 days will be studied, at various time points.

The experiments in this chapter are being conducted in response to the results of those carried out in chapter 4. Here, experiments suffered, primarily, from a lack of biological replication caused, in part, by the high demands of microarray experiments for relatively large amounts of RNA. Section 3 optimised an in-vitro transcription reaction to generate enough aRNA for approximately 50 microarray experiments from a single barrel cortex. This means that pooling of samples (as carried out in section 4) is no longer required and efficient replication can be incorporated into the experimental design.

The other major difference between the experiments conducted here is that they will be carried out using 15,000 gene microarray slides. Chapter 4 made use of a small-scale microarray that resulted in problems with normalisation (section 4.4). By using a large scale microarray a mathematical normalisation can be conducted, that will reduce the technical error that can cause false positives in the resulting data (Quackenbush, 2002). Special consideration will be given to the concept of experimental design, as it relates to microarray experiments, in order to find the most effective design for the analyses required.

5.2 EXPERIMENTAL DESIGN IN MICROARRAY ANALYSIS

The aim of this section is to find the most efficient design scheme for a series of microarray experiments that will analyse the changes in gene expression in the mouse barrel cortex, following vibrissal deprivation. Specifically, the impact on available

resources and the amenability of the data generated to statistical testing, using various designs will be investigated.

Microarray technology is typically used to analyse the differences in gene transcript levels in different RNA samples, a hypothetical example being untreated cells and cells treated with some drug. As the technology has developed, investigators have attempted to answer increasingly complex questions using microarray experiments eg multi-treatment effects (Oleksiak et al., 2002), time-course experiments with three or more time-points (Spellman et al., 1998) and multiple-factor experiments (Jin et al., 2001).

Using two-colour microarray technology, only two samples can be co-hybridised in a single experiment, thus in the hypothetical example above untreated cells would be co-hybridised on the same microarray slide. In complex experiments, with three or more treatments under comparison, how to pair samples on slides is not as self-evident. Without careful planning in the design stage statistical techniques that test the validity of the data may not be applicable and financial resources may have been wasted performing hybridisation that were not required. The field of experimental design, as related to two-colour microarray technology, has developed into an area of research in itself, focusing on maximising the design space of an experiment (Yang and Speed, 2002, Kerr and Churchill, 2001, Churchill, 2002, Dobbin and Simon, 2002).

5.2.1 REPLICATION

Replication is essential to any experiment that requires a level of statistical significance to be applied to the results. Replicating an experiment reduces the amount of uncertainty in the data by decreasing the effects of systematic and random variation, thereby increasing the accuracy of measurements generated. The microarray process is lengthy and expensive therefore early experimenters tended to neglect replication (Schena et al., 1995, DeRisi et al., 1996). However as increasingly complex questions, involving many different treatments are being asked in microarray experiments, replication is now considered essential to the process (Jin et al., 2001). Replication can be incorporated into a microarray experiment at a number of levels, the level classifying the replicate as either technical or biological.

An example of technical replication is the multiple spotting of slides. Here, as in the custom microarray used in section 4, each cDNA target is printed mutiple times on a microarray (Black and Doerge, 2004). As shown replicate spots can be used to assess the quality of a particular hybridisation (section 3.6). However, each replicate is subject to

the same systematic and random variation in the process and will be equally affected by them. This limits the independence of their measurements and consequently their use in statistical inference (Yang and Speed, 2002). It is recommended that replicated spots not be positioned adjacent to one another as local array effects eg bubbles, will reduce the variance of the replicates and underestimate the overall variance across the microarray (Churchill, 2002).

Data from replicate spots typically show a correlation in the region of 95%. If a full technical replication, in which aliqouts of the same probe are re-hybridised to a new slide is performed, this correlation falls to between 60-80%. If the technical replication is a dye swap, in which the experiment is replicated with the dye allocation reversed, this correlation falls further (Churchill, 2002, Kerr and Churchill, 2001). This suggests that between-slide variation is greater than intra-slide variation. Thus, the most efficient way of estimating the variance in a microarray experiment is by carrying out a technical replication, with dye swap. Though the results of such a replication will appear to suggest a loss of reproducibility they can be used to filter out spurious data and provide an idea of the noise in the data.

While technical replicates provide important information about the microarray process, they provide little information about the variation in the biological population from where samples were taken. Consequently in experiments where assumptions at the level of the organism are to made biological replication is essential ie if it is to be concluded that *gene x* is differentially regulated following chessboard deprivation, multiple animals must be suitably treated. Biological replication allows the variation of the population to be estimated and this provides the basis for declaring a gene's differential expression as statistically significant (Yang and Speed, 2002, Churchill, 2002).

The number of biological replicates to include in a microarray experiment is the source of much debate. In order to test a hypothesis traditional statistical techniques rely on power calculations which provide the probability that the null hypothesis under test can either be accepted or rejected in favour of an alternative (Speed, 2003). To carry out a power calculation the variance of individual measurements, the magnitude of the effect, the acceptable false positive rate and the probability of finding an effect that is equal to or greater than the specified effect size must all be stated eg it might be of interest to know how many replicates are required to have a 95% chance of finding genes that are differentially expressed at a 2- fold level. If the variance of the differentially expressed genes is known this could be calculated, however as the variance of ratios are known to

vary greatly across microarray hybridisations this part of the calculation is missing. Furthermore it is impossible to state in advance of an experiment what magnitude of fold-change is significant (Yang and Speed, 2002). Attempts have been made to answer the question of how many replicates to use through variability modelling (Zien et al., 2003) and direct testing (Lee et al., 2000). Lee et al found that at least 3 replicates significantly reduces the chances of false positives, while Zien et al (2003) suggest a minimum of 8 replicates, for experiments that generate high ratios.

It was decided to analyse the gene expression changes in the barrel cortex at five time-points; 0.4, 1, 2, 4 and 8 days deprivation. In order to cover these time points it was only physically possible to carry out three biological replications for each deprivation treatment; 3 all whiskers present, 3 all whiskers deprived and 3 chessboard deprived. However, this provides 6 degrees of freedom (9 independent replicates minus 3 treatments) with which to estimate variance (Churchill, 2002).

In addition to biological replication it was decided to incorporate one technical replication into the design, to further try and reduce the impact of variation. Technical replication cannot be incorporated at the multiple spot level as these hybridisations will be performed on microarrays constructed using every gene from the NIA 15K mouse set (section 2.15.3). This number of genes, using the printing technology available, completely fills a microarray slide with no room for replication. Therefore each hybridisation was repeated using the same samples but with the dye labels reversed, in a dye swap replication.

5.2.2 EXPERIMENTAL DESIGN COMPARISONS

Having decided on how to partition the replication in this series of experiments (section 5.2.1) the problem of how to pair the probes to maximise the efficiency of the technology needed to be considered. Ideally, every potential comparison should be carried out amongst the samples, however cost and time constraints make it an impossibility for this, and most studies. Specifically, this section will outline the possible designs for pairing experimental samples in these experiments. Consideration will be given to the implications of each design, with respect to the amount of input required to complete the experiment and the application of statistical testing facilitated by various designs.

Consider two RNA samples extracted from the barrel cortex of mice. One sample is from a barrel field whose contralateral whiskers were deprived in a chessboard pattern (CB) while the contralateral vibrissae of the second sample are undeprived (ALL). CB is labelled with Cy3, ALL is labelled Cy5 and the samples are co-hybridised to an NIA 15K

clone set microarray. Such an experiment is depicted graphically in figure 5.1, using a convention that is employed throughout microarray literature as an aid to visualising and describing the design of experiments (Churchill, 2002).

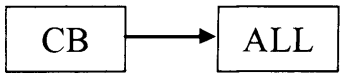


Figure 5.1 The graphical description of a two-colour microarray hybridisation. Boxes represent the RNA samples being contrasted. The arrow indicates both samples are being co-hybridised with the sample at the tail of the arrow (CB) labelled with Cy3 and the other sample (ALL) labelled Cy5.

In this series of experiments the overall aim is to identify genes that are differentially regulated only under vibrissal deprivation conditions that induce neuronal plasticity (CB). This necessitates the comparison of barrel cortices subjected to three distinct treatments; CB, ALL and all deprived (DEP). As microarray technology only permits the comparison of two treatments in a single hybridisation, the experimental design can be considered an incomplete block (Cochran and Cox, 1957) ie all distinct treatments are not present on a microarray. In section 4, this problem was dealt with by using additional experimental samples in a within-subject approach (figure 5.2a), that allowed all distinct treatments to be compared. An alternative approach would employ a between-subject design in which both sets of vibrissae in a single animal are subjected to the same treatment, either CB, ALL or DEP (see figure 5.2b).

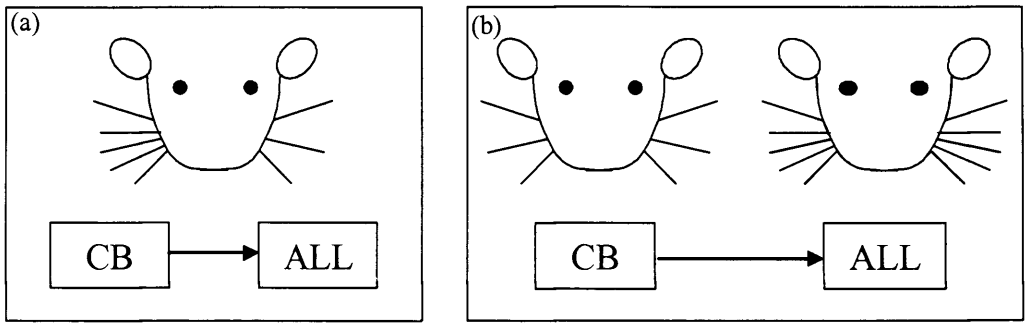


Figure 5.2 A comparison of the within-subject (a) and between subject (b) designs for a microarray hybridisation.

a) a single experimental unit (mouse) is subjected to two treatments (ALL and CB) while b) applies the two treatments to seperate units ie in a) each of the two sets of mystacial vibrissae are plucked using different treatments.

b) the same treatment is applied to both sets of vibrissae. Thus to compare CB and ALL barrel fields, design a) requires one animal while b) requires two.

In using the within-subject approach the design of the experiment is self-evident as the only comparison available is between the barrel cortices of the same animal. When using the between-subject method the pairing of samples is not so obvious. With 45 (five time-points, nine animals per time-point) experimental animals to compare, careful consideration must be given to the exact aims of this experiment. Every potential comparison is impossible to carry out, therefore resources should be focused on carrying out those hybridisations which provide the most useful data eg would pairing an 8 day DEP barrel cortex with a 1 day CB be as informative as the data from a hybridisation that compares 8 day DEP with 8 day CB? In order to compare various designs it may be prudent to consider first an individual time point in which three biological replicates, subjected to three treatments, are to be compared to one another with a technical replicate, or dye swap, of each hybridisation.

Three possible experimental designs that enable such comparisons are shown in figures 5.3a – c. Here the schematic representations outlined in figure 5.1 are used to describe the hybridisations required ie the biological replicates are placed in boxes and an arrow represents the fact that a hybridisation is carried out between the respective samples. The direction of the arrow describes the allocation of a particular dye, Cy3 or Cy5 in this case, to each sample. It is worth noting that opposite facing arrows between two samples represents the technical dye swap replication.

Figure 5.3a describes the design that must be used if within-subject comparisons are to be carried out ie the probes from the barrel fields of the same animal are co-hybridised. The design in figure 5.3b compares every sample to a reference RNA that may or may not be related to the experimental RNA. This reference design gives an insight into an important principle of two-colour microarrays; the fact that gene expression in any two samples can be contrasted as long as there is a comparison linking them (Churchill, 2002) Any RNA sample can be indirectly compared to another as long as they are both co-hybridised to a common RNA sample (the reference). Notice the use of the word 'indirectly' in the previous sentence. The use of a reference sample means that no hybridisations between any barrel cortex probe are actually carried out.

In contrast the loop design (figure 5.3c) actually combines the samples of interest directly in a hybridisation. Though it may appear daunting the loop design is actually a simple method of ensuring that all comparisons between treatments are carried out and, importantly, the labelling of dyes is balanced so that each sample is labelled four times, twice with Cy3, twice with Cy5.

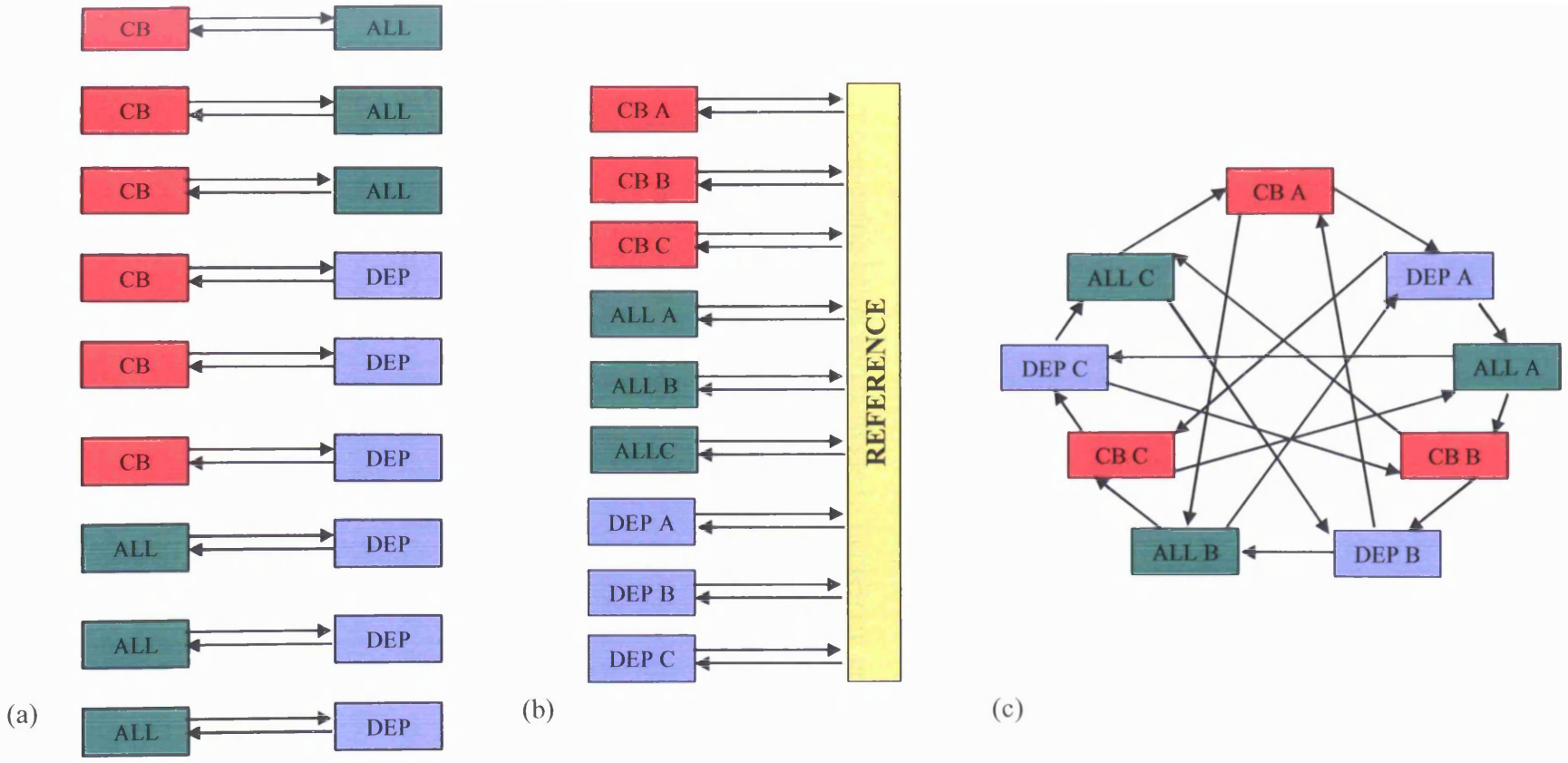


Figure 5.3 Possible experimental designs for comparing the gene expression of barrel cortices subjected to three whisker deprivations, or treatments (ALL-green, CB-red and DEP-blue) at 1 time point using microarray technology. In each design 3 independent biological replicates of each treatment are incorporated in addition to a technical replicate (dye swap). Each box represents an independent biological sample. Arrows indicate that the samples are compared in a two-colour hybridisation. The sample at the head of the arrow is labelled Cy3 and that at the base Cy5.

a) A within-subject design in which 2 treatments are applied to each animal. b) A between-subject design in which the same treatments are applied to a single animal. Each sample is compared to a common reference RNA sample allowing indirect comparisons to be made between each sample. c) A between-subject balanced 'loop' design in which treatments are compared through the direct comparisons of independent biological replicates. Note that the design is balanced with regards to the dye allocation so that each sample is labelled 4 times (2*Cy3 and 2*Cy5) and every treatment is compared to the others twice.

All three designs in figure 5.3 allow for differences in gene expression between CB, ALL and DEP samples to be carried out. Each design uses the same number of biological replicates (mice) and provide technical replication in the form of dye-swap hybridisations. Table 5.1 provides a breakdown of the resources required to carry out the hybridisations required to complete each design. The main difference, in terms of workload, between the three designs occurs when the within-subject comparison (figure 5.3 a) is compared to the reference and loop designs. Here double the number of RNA extraction and subsequent RNA amplifications are required. This is because each barrel field of the nine animals must be processed separately for the within-subject approach. In the between-subject approach both barrel fields from a single animal can be pooled and extracted / amplified together. Apart from this difference, which is not inconsiderable, the resource allocations are equal for each design.

When the experiment is scaled up to include five timepoints the doubling of RNA extractions / amplifications in the within-subject design is still evident. Though the multiple processing of samples would be used, the increase in processing time required, for converting barrel cortex tissue to fluorescent probe, would mean this design is highly undesirable. From table 5.2, the reference design requires far fewer labelling reactions and hybridisations than the loop design. This difference is quite substantial both financially and in terms of workload and occurs because extra hybridisations are required to perform the same number of comparisons when the loop design is used compared to the reference design. This means it is possible to compare the effect of various treatments over time eg the temporal effects of CB deprivation on gene expression in the barrel cortex over a period of time represented by the five time-points can be deduced. This is possible using the reference design as each sample is linked by comparison to the common reference sample. In order to achieve such comparisons using the loop design, additional loops are required that in effect link the loop shown in figure 5.3c with the loops that would be created at other time points. An example of one such additional time course loop such is shown in figure 5.4 and is the reason that extra hybridisations are required. The design of an entire series of experiments that use loops to carry out the same number of comparisons as the reference is shown in figure 5.5. This design is far more convoluted than the reference design shown in figure 5.6 and comparing the two diagrams gives an idea of the relative simplicity of reference designs.

1 Timepoint	Within-Subject Comparison	Reference Design	Loop Design
Experimental Units	9	9	9
RNA Extraction / Amplification	18 / 18	9 / 9 + 1 (Reference)	9 / 9
Labelling Reaction	36	36	36
Number of hybridisations	18	18	18

Table 5.1 Resource allocations for possible experimental designs for comparing the gene expression of barrel cortices subjected to three whisker deprivations, or treatments (ALL, CB and DEP) at 1 time point using microarray technology. In each case 3 independent biological replicates are carried out in addition to a technical replicate involving a dye swap. Each major step involved in the process is included to provide an idea of the practical and financial implications of using each design. Note how the within-subject design requires twice the number of RNA extractions when compared to the loop and reference design. Also of note is that resources must be allocated to amplification of the reference sample in addition to its labelling.

5 Timepoints	Within-Subject Comparison	Reference Design	Loop Design
Experimental Units	45	45	45
RNA Extraction / Amplification	90 / 90	45 / 45 + 10 (Reference)	45 / 45
Labelling Reaction	180	180	264
Number of hybridisations	90	90	132

Table 5.2 Resource allocations as described for table 5.1 when the number of time points used in the experiment is increased to 5. The within-subject design requires twice the number of RNA extractions and amplifications compared to reference and loop. Now more labelling reactions are required for the loop than the reference due to the extra hybridisations required to compare each treatment in the time course side loops (see figures 5.5 and 5.6).

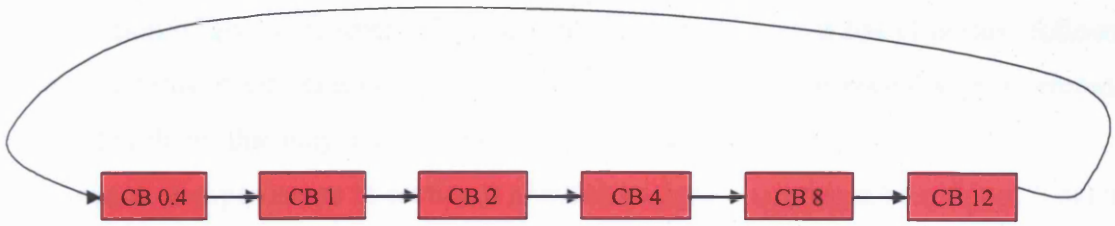


Figure 5.4 Experimental design for a time course analysis of gene expression in the mouse barrel cortex. Animals are subjected to chessboard deprivation for a set period of time (indicated by numbers in box). Each sample is arranged in hybridisations in a loop design (for convention see figure 5.2) that balances all samples and dye allocations. This example uses one biological replicate for each time point. Replication is achieved simply by repeating the experiment but reversing the dye order ie arrows will be reversed. This design is equally applicable to any of the treatments (ALL, CB or DEP) in order to monitor the temporal effects of whisker deprivation.

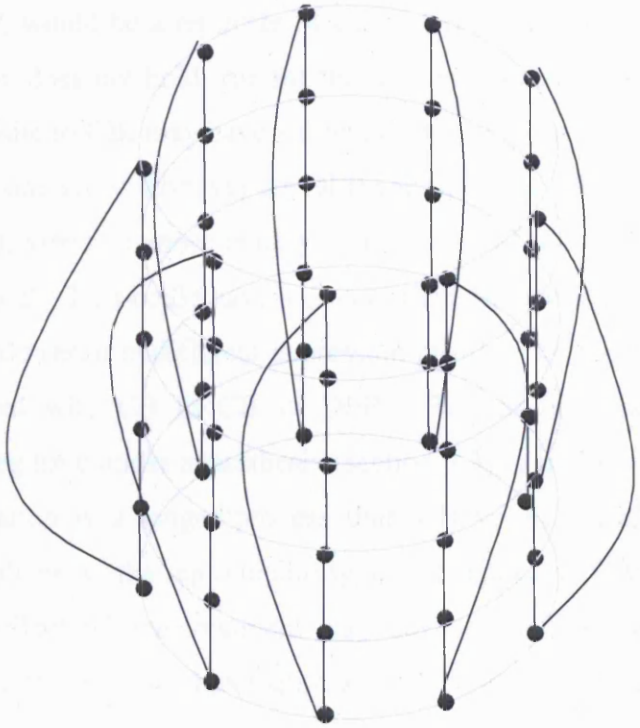


Figure 5.5 A schematic illustration of a complete loop design for analysing the effects of gene expression in the barrel cortex following whisker deprivation. In addition to the analysis of the effects of three treatments at single time –points on gene expression, the temporal effect of whisker deprivation can be discovered eg a time course experiment can be conducted in which the effects of CB deprivation with time on gene expression can be deduced. The loop required for the analysis of a single time point, as depicted in figure 5.3 c, is shown in grey. Each time point is connected by additional loops shown for one replicate in figure 5.4 and shown as black lines here. Black circles indicate experimental samples.

This section has highlighted three potential experimental designs that would allow the identification of genes, differentially regulated within the mouse barrel cortex, following various patterns of vibrissal deprivation. Both the loop and reference design overcome a potential problem that may occur using the within-subject design.

In the series of experiments in section 4 a within-sample comparison (see figure 5.2a) was employed in which two separate treatments, or whisker deprivation patterns, were applied to each animal (experimental unit). Therefore, three hybridisations were required to compare the effects, on gene expression in the mouse barrel cortex, of all three treatments. The experimental aim was to find genes regulated under CB conditions by comparing, CB with DEP, and, CB with ALL, and identifying differentially expressed genes that were common in both comparisons.

The assumption was made in chapter 4 that the gene expression profile of CB treatment, in CB vs DEP, would be a replicate of CB in CB vs ALL. However, it is possible that this assumption does not hold true for the following reason; the treatment applied to the vibrissae opposite to CB, may have a direct effect on the animal's use of the CB deprived vibrissae eg if one set of vibrissae are DEP and the other CB, then the animal may use its remaining, CB, vibrissae more readily than an animal in which one set of vibrissae are ALL and one CB. This could have a direct effect on activity levels within the CB barrel cortex that could result in different expression profiles, for the CB barrel cortex in CB vs ALL, compared with CB in CB vs DEP. Furthermore, the effect of experimental manipulation eg time under anaesthesia (section 2.3) is poorly controlled for. The CB vs DEP manipulation is a longer process than CB vs ALL and this could increase the variability, or decrease the reproducibility, of the replicated CB gene expression profiles. In short, the effect of one treatment may have a confounding effect on the opposite treatment using the within-subject approach. This confounding may not occur and both CB barrel fields can be considered as replicates, however, without further testing there is no way of showing this. Chapter 3 highlighted the variability inherent in microarray experiments. Indeed even a self-self hybridisation (section 3.6.1 for discussion) will show some variation from the expected results. It is therefore imperative that no further potential sources of variation are incorporated into the design.

The loop and reference design offer an alternative approach, employing a between-subject design in which both sets of vibrissae are subjected to the same treatment, either CB, ALL or DEP (see figure 5.2 a and b). In this case the confounding effects described above are controlled for and all three treatments could be compared effectively freeing

the design from the constraints imposed by the within-subject comparison. For this reason the within-subject design can be ruled as a potential candidate design for this series of experiments.

The within-subject approach is also far more labour intensive requiring twice the amount of sample processing than either the loop or reference designs. This imposes increased pressures on the experimental resources available, as does the extra labelling / hybridisations required for the loop compared to the reference designs (table 5.2). Thus, it appears from a financial and labour perspective, that the reference design is the most economical of the three.

However, it should be pointed out that the reference design and loop design (figures 5.3b and c respectively) compare the levels of gene expression in two samples in different ways. The loop design is a direct comparison of gene expression in two treated samples ie the ratio of a gene in one sample compared to another is estimated using:

log_2 *(Cy3/Cy5) eg* log_2 *(CB/DEP)*

The reference design estimates this measure indirectly by:

$log_2 (CB/DEP) = log_2 (CB/REF) - log_2 (DEP/REF)$ Eqn. 5.1

The implication of this is that the indirect (reference) measurement has twice the variance, or half the accuracy, of the direct (loop) design (Yang and Speed, 2002), as two hybridisations are required to obtain the measurement. Every hybridisation measurement has a degree of variance associated with it, thus, if two hybridisations are required then the variance in the measurement is doubled.

Loop designs were first described by Kerr and Churchill who noted that half of all data collected in a reference design is on a sample of no interest. They proposed that a loop design makes more efficient use of the design space available and through proper design the loop overcomes the incomplete block design (Cochran and Cox, 1957) that occurs when more than two treatments are compared in two-colour microarray. Using a loop design any pair of treatments occur together on a microarray the same number of times, the dyes allocated to each sample are balanced ie the design is *A-optimal* (Kerr and Churchill, 2001).

Kerr and Churchill have based their ideas on statistical techniques and experimental design principles developed for studying the yield of different varieties of crops (Fisher, 1951). Recognizing similarities between the crop yield and microarray experiments, Kerr and Churchill realised that Fisher's statistical rules could be applied to data generated in microarray experiments, but only if the experiment was correctly designed ie a loop design was used.

Fisher realised that proper statistical comparisons of crop yields could only be made if the crops were grown on the same plot. If two varieties of crop are grown on different plots any difference in yield could be down to a number of factors such as sunlight, climate or soil fertility. The analogy in microarray experiments is that treatments (or deprivation patterns) can only be compared if they are co-hybridised on the same array (or a reference design is used), otherwise the measurements are confounded with slide-slide variability Churchill, 2002). The influence of these confounding factors prevents any statistical significance being attached to the effect of variety unless the varieties are grown on the same plot.

In addition, Fisher realised that growing the same variety on a number of plots allows the effect of each plot to be estimated in addition to the effect of variety leading to the development of ANOVA (Analysis of Variance), a statistical test used to compare more than two varieties (or treatments). Using these ideas Kerr at al have developed a model that estimates the effects of confounding factors that contribute to error in microarray data (Kerr et al., 2000). The model is presented in equation 5.2:

$$log\ (y_{ijkg}) = \mu + A_i + D_j + G_g + (AD)_{ij} + (AG)_{ig} + (VG)_{kg} + (DG)_{jg} + \epsilon_{ijkg} \quad \text{Eqn 5.2.}$$

$log\ (y_{ijkg})$ is the signal for gene $_g$ on array $_i$ for dye $_j$ and treatment $_k$.

μ represents the average signal across all arrays.

A_i is the effect of Array $_i$

D_j is the effect of dye $_j$

G_g is the effect of gene $_g$

$(AD)_{ij}$ is a term for the combined interaction of Array $_i$ and Dye $_j$

$(AG)_{ig}$ is a term for the combined interaction of Array $_i$ and gene $_g$

$(VG)_{kg}$ is a term for the combined interaction of treatment $_k$ and gene $_g$

$(DG)_{jg}$ is a term for the combined interaction of dye $_j$ and gene $_g$

ϵ_{ijkg} is the effect of random noise

By using direct comparisons in a balanced loop design it is possible to estimate the contribution of these factors to measured expression levels. The factor of most interest is the $(VG)_{kg}$ term which provides a measure of the combined effect of treatment and gene. As in standard ANOVA models the effects are used to calculate residuals which estimate the effect of the noise term ϵ_{ijkg}. If the distribution of the residuals is non-random then all sources of variation have not been captured indicating an inadequate experimental design. Importantly, the variability of the factors can be subtracted from the overall variability such that the $(VG)_{kg}$ term is estimated with greater accuracy, allowing more efficient testing of hypotheses regarding differential gene expression (Kerr et al., 2001).

An implicit understanding in this model is that expression values are not converted to ratios. When using the reference design to compare across samples (Eqn 5.1) without converting to ratios the observed measurements generated by the reference design are confounded with slide-slide variability, therefore the ANOVA approach cannot be applied.

Kerr, Churchill, and others argue that measures taken on a reference sample are unnecessary and waste the design space of the experiment (Kerr and Churchill, 2001, Dobbin and Simon, 2002). In addition the effect of the dye gene ($(DG)_{jg}$ eqn 5.2) interaction is not accounted for in the reference design. This was true for many, early microarray experiments which used a reference design. Typically all hybridisations were run forwards ie all reference samples were labelled with Cy5 (eg Boldrick et al., 2002) so that dye effects and variety effects are completely confounded. However it has been shown that including a 'few' reverse arrays into the reference design, if comparison with the reference is of primary interest, can reduce the impact of this bias (Dobbin et al., 2003). In situations, like here, where comparison with the reference is secondary all arrays should be run forwards.

Essentially the argument against using a reference design amounts to the fact that the same information, with less error, can be obtained using fewer hybridisations by incorporating samples in a loop design. This may be true when small numbers of samples are to be compared, however, as can be seen in table 5.2 the reference design uses fewer resources than a loop for all potential comparisons in this series of experiments. Despite the fact that the reference design estimates variance with lower efficiency than the loop and that the ANOVA model presented (Eqn 5.2) is inapplicable to a reference design does not mean that the reference design is not amenable to statistical analysis (Dobbin and Simon, 2002).

Finally one important failing of the loop design must be taken into account ie the fact that if two or more hybridisations fail the entire ANOVA model breaks down (Simon et al., 2002). From past experience (section 4) it is not unknown for hybridisations to completely fail. If this occurred when a loop design was being used extra hybridisations would be required before any data could be extracted. On the other hand the reference design is relatively stable to experimental failure and data can be extracted and interpreted from comparisons without carrying out every hybridisation outlined in the design (Simon et al., 2002). Thus for its robustness and economic saving the reference design will be used for this series of experiments.

In choosing a suitable reference RNA sample thought was given to the use of a commercially available universal RNA, supplied by companies such as Stratagene® (CA,USA). The aim of this reference is to enable comparisons to be made between experiments conducted, for example, in different labs. Universal references are designed to incorporate as many genes as possible, at equal levels, in order to 'light up' the majority of spots on any microarray thus maximising the useable number of data points in a hybridisation. While a good idea in theory it is perhaps best that the reference RNA chosen be related to the RNA samples under investigation otherwise extreme differences in ratios could occur for many genes which would effect the process of normalisation (Quackenbush, 2002). Therefore a reference sample will be generated through the pooling of experimental samples (see section 2.8.8).

5.3 MIAME COMPLIANCE

This section provides the reader with a review of some of the key concepts used in a series of microarray experiments, designed to analyse gene expression in the mouse barrel cortex, following various forms of whisker deprivation. In addition to being a useful reminder, this section follows the guidelines set out in the Minimum Information About a Microarray Experiment (MIAME) standards protocol which aims to provide a framework for the reporting of microarray data (Brazma et al., 2001). MIAME was proposed by the Microarray Gene Expression Data Society (MGED, http://www.mged.org/index.html), an international organisation of biologists, computer scientists and data analysts who recognised a need to generate a standardised platform by which, microarray data could be stored, accessed and interpreted by the scientific community at large.

The development of MIAME came through the recognition that while the vast amounts of data generated through large-scale genome sequencing projects data can be readily accessed and interpreted by the scientific community through the use of standard formats, software tools and databases, the data from microarray experiments is largely inaccessible. This is partly to do with the fact that there is no standard technique used to conduct a microarray experiment, generate the data, transform and interpret it eg different laboratories use different microarray slides, labelling techniques, normalisation strategies etc. Also, microarray technology is not a quantitative subject ie it does not provide absolute levels of gene expression using units of measure. The data is usually reduced to a relative measure which means that the comparison of gene expression data generated by different groups is difficult.

It is clear that, while standard protocols for conducting, analysing and interpreting microarray experiments may be a long way off, by reporting the experiment in a standard way the interpretation, analysis and verification of data by other investigators will be simplified. Essentially MIAME is a checklist that should be carried out prior to reporting a microarray experiment. By satisfying all of the criteria an investigator will provide any readers with all the information required for the complete analysis and interpretation of the data. Thus the series of experiments conducted here will be described using the guidelines set out in MIAME.

5.3.1 EXPERIMENTAL DESIGN

This experiment was designed to analyse the effects of vibrissal deprivation on gene expression, in the barrel cortex of mice, using microarray hybridisations in which experimental samples were co-hybridised with a reference sample. All hybridisations performed are depicted in figure 5.6, using the standard convention employed by Churchill (Churchill, 2002), in which samples are placed in boxes and a direct hybridisation between samples is indicated by an arrow. The sample at the tail of the arrow is labelled with Cy5 fluorophore while that at the head is labelled Cy3. Each box represents an individual biological unit subjected to one of 15 particular treatments.

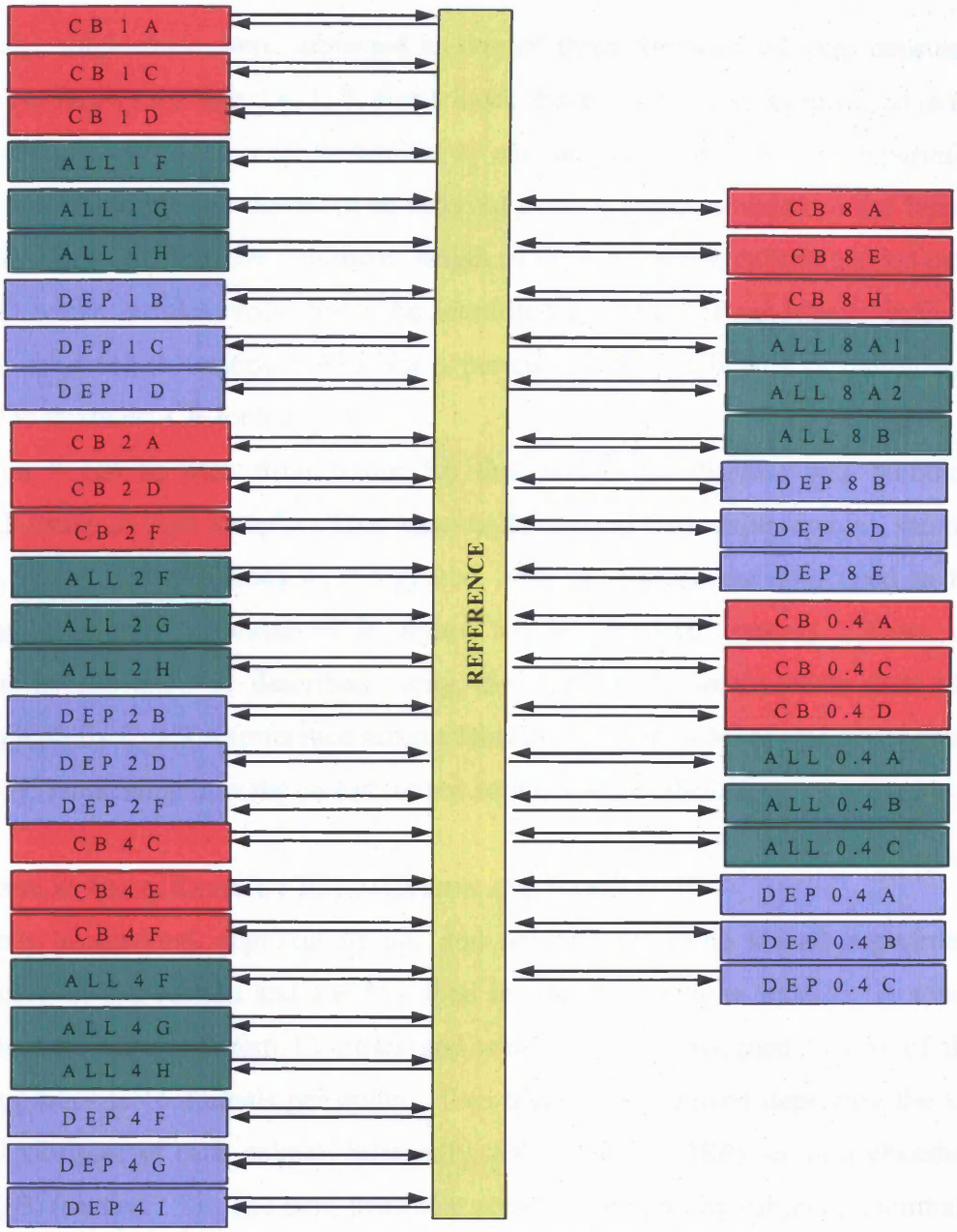


Figure 5.6 The design of a series of experiments studying the effects of vibrissal deprivation on gene expression, in the barrel cortex of mice, using microarray technology. Each box represents an individual biological sample. Arrows represent the fact that samples are co-hybridised on a microarray slide. The direction of the arrow describes the dye used to label each sample. The sample at the tail of the arrow is labelled with Cy5 fluorophore while that at the head is labelled Cy3. In this case every experimental sample is aRNA generated from the pooled barrel cortices of a single animal. Each sample is assigned a code that describes the treatment applied to both sets of vibrissae which were either chessboard deprived (CB, red), every vibrissae deprived (DEP, blue) or undeprived (ALL, green). Deprivations were carried out for periods of time indicated by the number in each sample code (either 0.4, 1, 2, 4 or 8 days). A single letter describes individual samples from within each treatment group. Thus it can be seen that all samples were co-hybridised with a reference sample (see text) and replications of each hybridisation were carried out in which the dyes were swapped.

Specifically, single mice were subjected to one of three forms of whisker deprivation (ALL, CB or DEP) for either 0.4, 1, 2, 4 or 8 days. Each treatment was replicated in three independent animals, which means a total of 45 animals were used for these experiments. Each individual sample can therefore be referred to by a code, provided in the boxes in figure 5.9, which describes the treatment, length of time the treatment was carried out for and an ID which enables replicates to be identified eg ALL 1 F represents individual animal F, subjected to treatment ALL for a period of 1 day; CB 8 A is individual A, subjected to treatment CB for 8 days etc.

In addition it can be seen from figure 5.6 that every hybridisation was technically replicated using a 'dye swap'. This essentially means that experimental sample / reference sample hybridisations were repeated, with the fluorescent dyes used to label each sample reversed (represented in figure 5.6 by reversed arrows). Thus, each individual hybridisation is described using the individual sample code and where necessary a prefix Cy3 Ref (reference sample labelled Cy3) or Cy5 Ref (reference sample labelled Cy5) indicating how the co-hybridised samples were labelled.

5.3.2 SAMPLES USED, EXTRACT PREPARATION AND LABELLING

The genetic background, approximate age and sex was the same for all experimental animals used in this section and are described in detail in sections 2.1-2.2. A total of 45animals were obtained from 13 litters and were randomly assigned to one of three treatment groups ie 15 animals per group. Two treatments involved depriving the large mystacial vibrissae of each animal, bilaterally, either totally (DEP), or in a chessboard pattern (CB) (section 2.3). The third treatment acted as a control by subjecting animals to the effect of anaesthesia without removing any vibrissae (ALL). Deprivations were carried out over a period of 8 days, during which time animals were routinely checked to ensure no vibrissal re-growth had occurred. Checking was carried out on alternate days from the start of deprivation and involved anaesthetising the animals, visual inspection and the removal of any re-grown vibrissae. Note that 'ALL' control animals were subjected to anaesthesia and put through a 'sham' checking procedure.

At specific timepoints, namely 0.4, 1, 2, 4 and 8 days following the start of treatment 3 animals were randomly removed from each treatment group for sample processing. Following sacrifice, brains were removed and both barrel cortices from each animal were dissected out and pooled (section 2.6). Total RNA was extracted from each pair of barrel cortices (section 2.8.2) and was subjected to quality control measures that ensured the

purity and integrity of each sample (section 3.2.2). Furthermore, the area of cortex sampled was visualised to ensure efficient sampling had occurred (section 2.7). Aliquots were taken from each of the 54 samples and pooled to form a reference batch of total RNA. Individual samples were put through an in-vitro transcription reaction to generate aRNA (section 2.8.6 and section 3.2.4). In addition aRNA was generated (approximately 400μg) from the reference total RNA, which would provide enough sample for 400 hybridisations. Again aRNA was subjected to quality control measures as described (section 3.2.4).

Experimental aRNA samples were prepared for hybridisations by labelling with either Cy3 or Cy5 fluorophore carried out using the CyScribe™ Direct Labelling kit from Amersham Bioscience, UK (section 2.12.6). To minimise variability a requirement of the reference sample was that labelled aliquots used in hybridisations were as homogenous as possible. To this end enough reference aRNA was labelled with Cy3 and Cy5 dye (section 2.12.6) to allow for 75 hybridisations each. The labelled reference was then pooled and split into 75 equal aliquots which were stored until required. By pooling, variations in individual labelling reactions were countered and the homogeneity of the reference probe was increased. Also, more labelled probe was generated than required to avoid having to re-label the reference if any hybridisations failed.

5.3.3 HYBRIDISATION PROCEDURES AND PARAMETERS

A thorough description of these procedures are provided in sections 2.16.

5.3.4 MICROARRAY SLIDE DESIGN

Microarrays used in this section were created by printing cDNA, generated from each bacterial clone in the NIA 15K clone set (section 3.3.1, http://www.cf.ac.uk/biosi/research/molbiol/index.html), onto CMT-GAPS™ (Corning, NY, USA) coated slides (section 2.15.3). The clone set is freely available to academic institutions through licensed distributors, in this case Cardiff University School of Bioscience Molecular Biology Support Unit:
(http://www.cf.ac.uk/biosi/research/molbiol/index.html)

Each of the approximately 15000 cDNA targets were singly spotted onto the microarrays in a particular layout. Appendix 4 ties in each spot on the microarray to an individual clone ID used by the NIA (column A). The co-ordinates of each spot are provided by its meta grid location and subgrid location (figure 3.16). Essentially when the microarrays

are printed the spots are laid down in blocks of size 22 x 22. There are a total of 48 blocks laid down in a so-called meta-grid. Thus four co-ordinates describe the position of each spot, the Meta Row, Meta Column, Row and Column (columns B - E respectively, appendix 4). In order to obtain further information about each clone, the NIA ID can be used to search the NIA database (http://lgsun.grc.nia.nih.gov/cDNA/15k.html) from where links to various information eg clone sequence are provided. The nucleotide database at the NCBI is also a useful tool which allows the sequence of any NIA 15K clone to be traced simply by entering the NIA ID (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=nucleotide).

In order to generate full length sequence information for each clone, during creation of the set each clone was sequenced from both its 3' and 5' ends. The sequence derived from either end can be obtained by using the additional clone IDs provided in columns G and J of appendix 4. These IDs can be used in the same way as those in column A to gain sequence information about the clone.

Importantly for each sequence (3' or 5'), the creators of the clone set have carried out BLAST searches using their derived sequences to identify clones with significant homology to sequences deposited in the NCBI database. Therefore, where possible, NCBI Genbank IDs have been provided, in columns H and K of appendix 4, for clones with such homologies. Many of these Genbank IDs correspond to sequences of known genes therefore putative gene IDs for each clone are provided, when available, in columns I and L of appendix 4. Columns N – S provide additional links to various databases which may prove useful, for finding out additional information about the gene in question that a particular clone shows homology to. In order the databases can be found using the following URLs:

Column N NCBI Unigene - http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene
Column O Gene Card Database - http://bioinformatics.weizmann.ac.il/cards/
Column P LocusLink - http://www.ncbi.nlm.nih.gov/projects/LocusLink/
Column Q The Jackson Lab- http://www.jax.org/
Column R TIGR Mouse Gene Index ID - http://www.tigr.org/

In order to use these databases the relevant ID should simply be entered into the search tool available at each site. It should be stressed that the information in columns G – S is

often updated and though correct at the time of writing additional information may be available at later dates eg more clones may be annotated.

Quality control procedures monitored the printing process. Here printed microarrays were randomly selected from print batches and tested for quality of printing (section 2.15.5). It has been shown that the quality of data obtained from microarray experiments can vary considerably and this variation can be attributed, in part, to the microarrays themselves (section 4.3). Although the variation cannot be claimed to be a result of deficiencies of the glass slides (eg contamination) or the print process its effect can be reduced by randomisation (Churchill, 2002). Therefore, microarrays for this series of experiments were printed in 3 batches of 50 and each hybridisation was allocated a slide by randomly picking one from the set of 150.

5.3.5 MEASUREMENT DATA AND SPECIFICATIONS

All microarray images were analysed using Imagene™ (Biodiscovery, Inc. CA, USA) microarray analysis software (section 3.6). The .txt output files from Imagene™ provide all information from this quantitation stage including; raw quantified intensities, spot diameter, spot flags etc. This data is provided in two folders in appendix 5, named Cy3 Ref or Cy5 Ref after the fluorophore used to label the reference sample. For each hybridisation two .txt files are provided, named using the experiment sample code (section 5.3.1) followed by the suffix 'A' or 'B'. Files with suffix 'A' represent the Cy3 channel of a particular experiment while those in 'B' represent the Cy5 channel. The identity of each spot is provided by either its four figure microarray co-ordinate or its NIA ID (section 5.3.4). The data can be imported directly into the GeneSight® software suite or, any microarray analysis software package that allows for .txt files.

These two .txt files were indeed imported into the GeneSight® software suite (BioDiscovery, Inc.,CA, section 3.6), which combined and paired the data from each channel. The data from each hybridisation was collated into two .txt files provided in appendix 6. One file provides the raw data from the Cy3 Ref experiments (reference sample is labelled Cy3) while the other is the data from Cy5 Ref experiments (reference sample is labelled Cy5). Again the NIA ID of each spot is given in column 1 while the remaining columns provide the background corrected median Cy3 intensity and background corrected Cy5 intensity for each spot in a particular hybridisation (for reasons why the median intensity was chosen see section 3.6.5). The particular hybridisation providing the data can be obtained from the column headers which use the

code for each experimental sample as an identifier followed by the suffix Cy3 or Cy5 indicating which channel the data was obtained from. Thus, for each microarray hybridisation the reference signal and experimental signals can be deduced from the name of the file containing the relevant data ie either Cy3 Ref or Cy5 Ref.
The 'raw data' provided in appendix 6 was subjected to further transformations using the GeneSight® software suite (BioDiscovery, Inc.,CA). In order the transformations were:

a) the removal of spots not significantly greater than background (section 3.7.2).

b) $\log_2$ transformation of the Cy3 / Cy5 quantified spot intensities (section 3.7.1).

c) Lowess normalisation (section 3.7.5)

The transformed data is provided in appendix 7, again in two .txt files one termed Cy3 Ref and one Cy5 Ref. Like the raw data, each spot on the microarrays are identified by their NIA clone ID (provided in column 1). The other columns provide the transformed quantified data for each channel (Cy3 or Cy5) with headers that describe the particular hybridisation that generated the data. Again using the file name the labelling of samples in each hybridisation can be deduced ie Cy3 Ref reference was labelled Cy3, Cy5 Ref reference was labelled Cy5.
Finally in appendix 8 the data used to identify significantly differentially expressed genes is provided. Again two .txt files are provided one named Cy3 Ref, one Cy5 Ref. Spots are identified using the NIA clone ID in column 1. In the remaining columns is the $\log_2$ ratio of the transformed experimental signal intensity to the transformed reference signal intensity (note: to avoid confusion $\log_2$ Cy3/Cy5 was not used). By using the ratio of experiment vs control, upregulated genes in a certain sample will have a higher positive ratio than other samples whilst downregulated genes will be more negative. This avoids confusion when comparing dye swap data as the ratios represent the same measure ie they are not inverted). The particular hybridisation which generated the data is provided by the code in the column header (section 5.3.1 for interpretation) and the name of each file provides the direction of sampling labelling.

5.4 MICROARRAY RESULTS

RI plots were generated for each hybridisation shown in figure 5.6. As examples the transformed datasets from each of the 18 hybridisations carried out on 1 day deprived samples are provided in figure 5.7 and 5.8. All RI plots showed the expected shape of microarray data that is not affected by systematic variations caused by the microarray process ie the data points show no significant curvature that would be indicative of a dependence of the $\log_2$ (Experiment/Reference) ratio on the intensity of the spot. The clustering of spots around R=0 is to be expected as the majority of spots in each hybridisation are not expected to show any differential expression. The lack of curvature in the data suggests that the LOWESS normalisation has removed any systematic dependence of the $\log_2$ (experiment intensity/reference intensity) on the signal intensity (section 3.7.5).

It was of interest to view the distribution of the $\log_2$ (Experiment/Reference) ratios of each hybridisation therefore distribution histograms were constructed for the data from each microarray hybridisation (figure 5.9). As discussed, one purpose of applying the log transformation to microarray data is to make microarray data approximate a normal distribution (section 3.7.1). The bell shape of the plots in figure 5.9 is characteristic of a normal or Gaussian distribution and suggests the log transformation has had the desired effect. Whether or not the distribution of the $\log_2$ (Exp/Ref) ratios are normally distributed has implications on the statistical testing applied to the data. Each distribution was tested using the Ryan-Joiner (R-J) test for normality (section 2.20.5). The results of the test are shown graphically in figure 5.10 and are summarised in table 5.3.

The R-J test essentially analyses the correlation between the observed data (in this case the distribution of $\log_2$ (Exp/Ref) ratios of each microarray) and the normal probabilities of the data in each bin of the histogram. Using the mean and standard deviation of the distributions of the observed $\log_2$ (Exp/Ref) ratios a hypothetical, normally distributed model is derived. When the data in each bin of this hypothetical dataset are plotted against their normal probabilities a straight line (blue in figure 5.10) results with a correlation co-efficient (or R-J value in table 5.3).

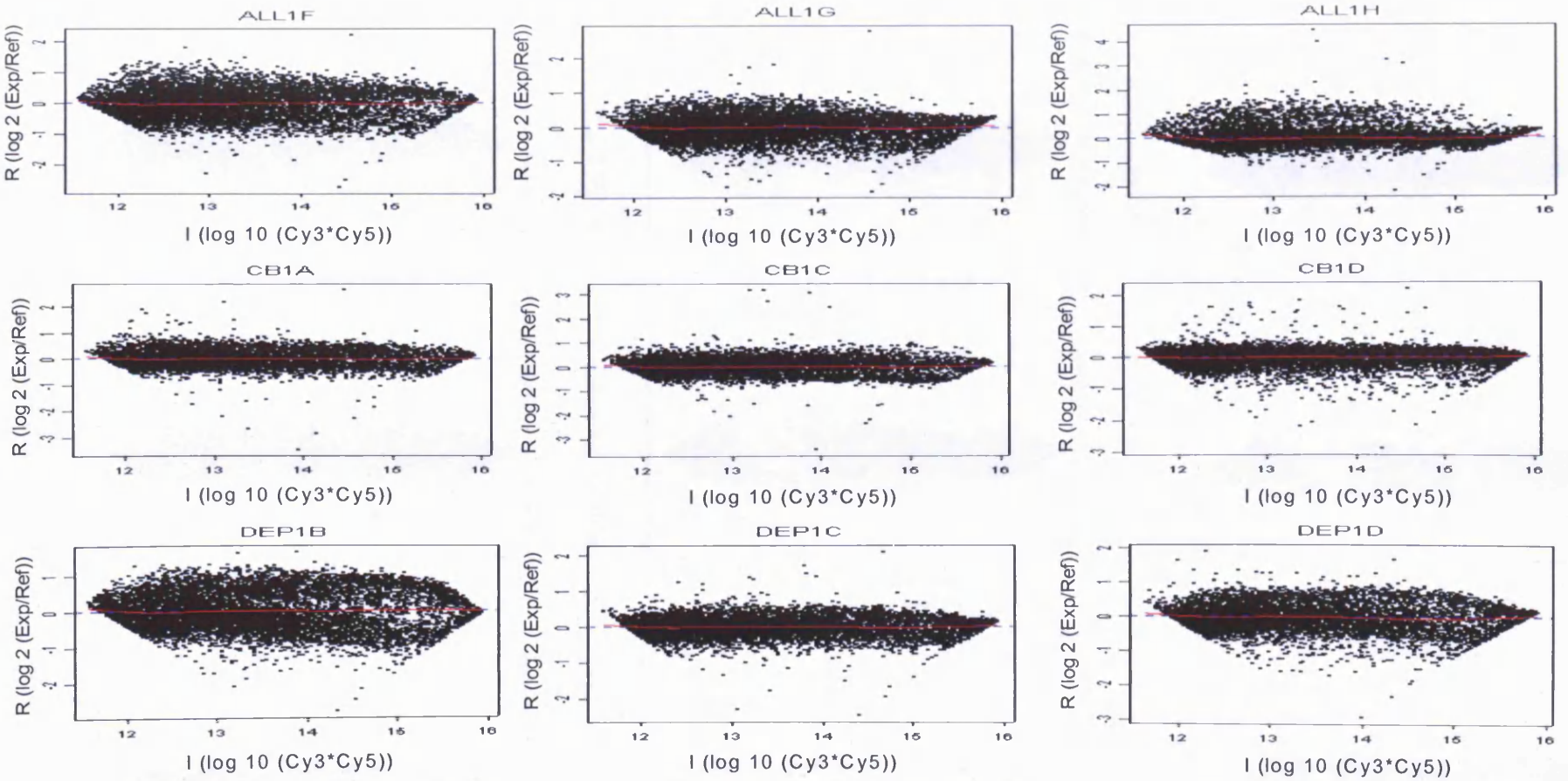


Figure 5.7 RI plots of the data from microarray hybridisations designed to analyse the effects of vibrissal deprivation, for a period of 1 day, on gene expression in the mouse barrel cortex. Suitably deprived animals were co-hybridised with a reference sample. The title of each plot represents the sample code used to describe each individual animal (see section 5.3.1 for further details). The RI plots represent the log2 (Experiment Intensity / Reference Intensity), y- axis, plotted as a function of the log10 (Experiment Intensity / Reference Intensity), x-axis, for each spot on the array, and are used to determine whether or not the magnitude of log ratios are dependent on overall spot intensities. The lack of a clear upward/downward trend in the data clouds indicate that transformations applied to the raw data have removed any systematic effects that may have been present. This is further emphasised by the fact that the red (lowess curve fit) line lies approximately parallel to R=0, suggesting that the majority of genes do not show differential expression.

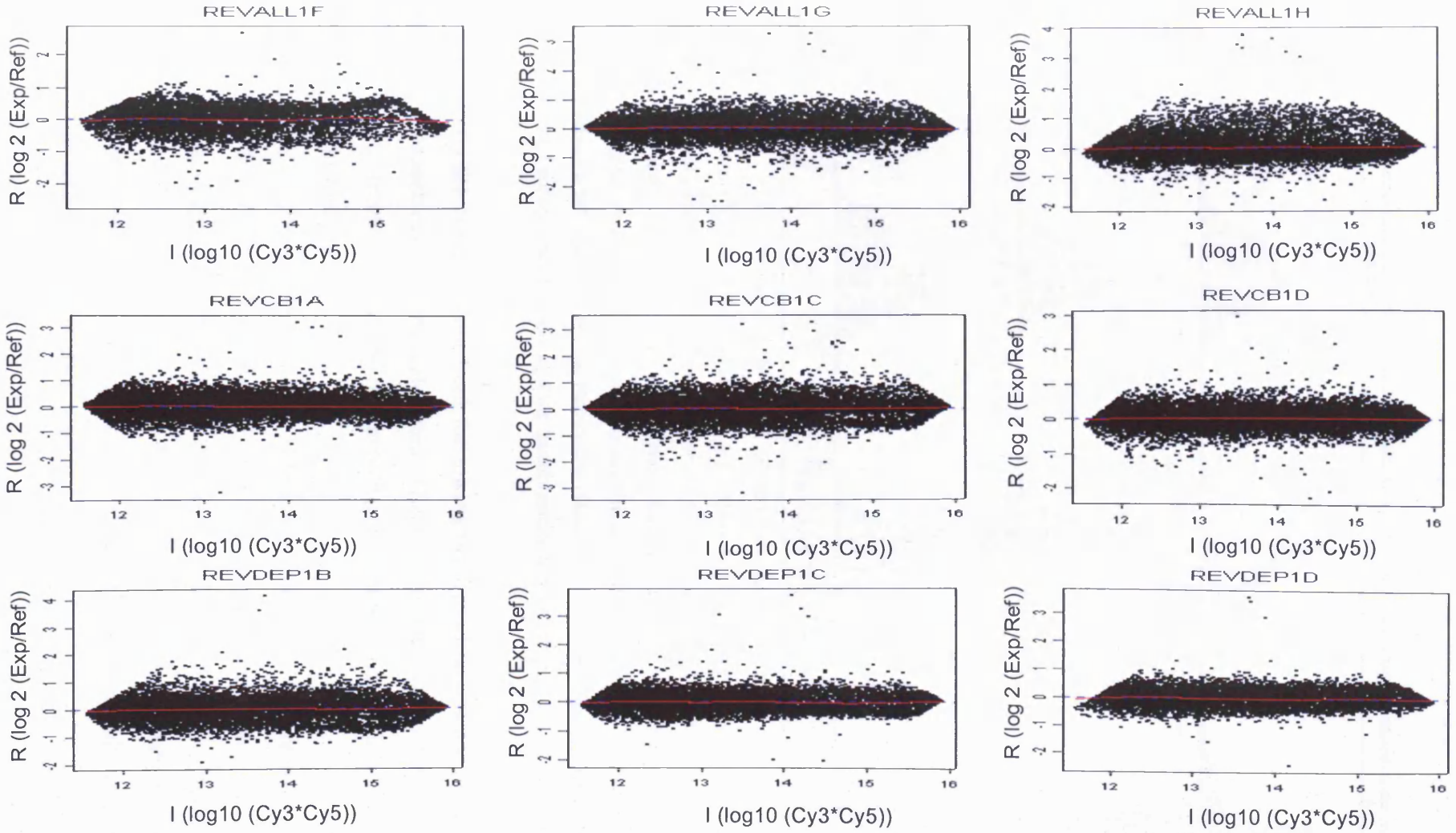


Figure 5.8 RI Plots of the Data from Dye Swap Hybridisations Analysing the Effects of 1 days Vibrissal Deprivation on Gene Expression in the Mouse Barrel Cortex. The identity of each hybridisation is given in the plot title. On each plot R ($\log_2$ (Experiment/Reference) or $\log_2$ (Cy3/Cy5)) is plotted against I ($\log_{10}$(Cy3*Cy5 intensity) for each spot on the NIA 15K microarray.

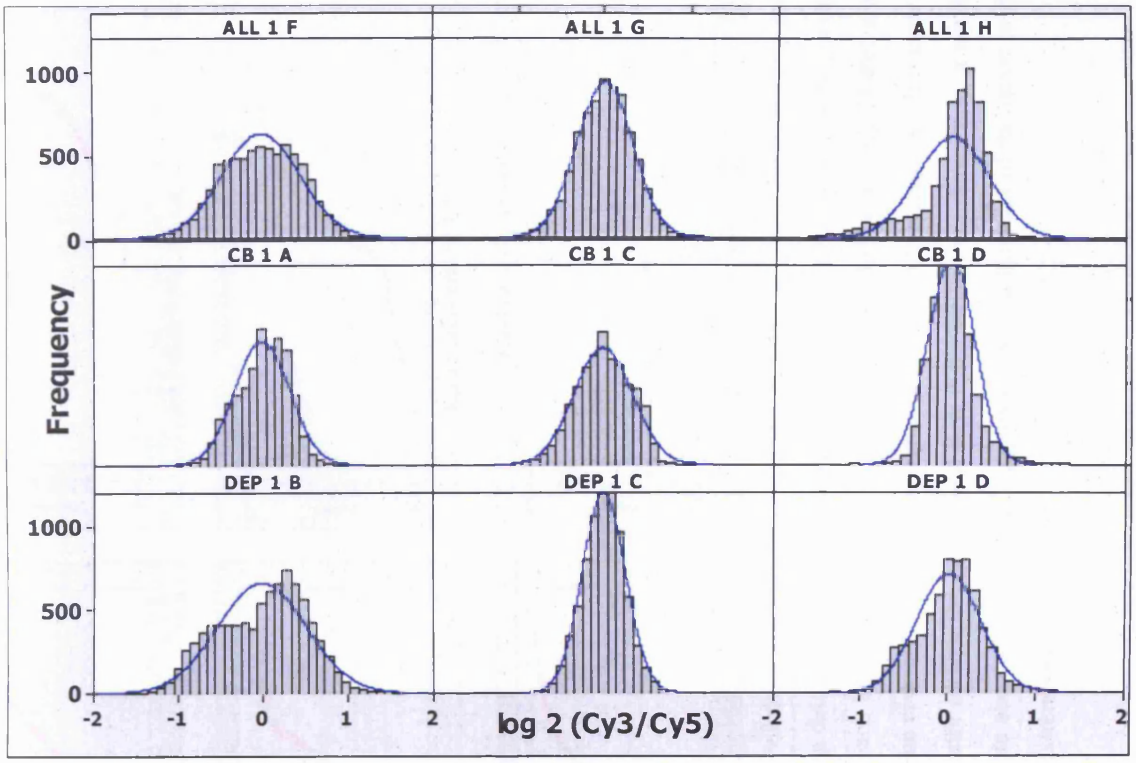


Figure 5.9 The distribution of the log 2 (Experiment/ Reference) ratios for microarray experiments designed to analyse the effects of vibrissal deprivation, for a period of 1 day, on gene expression in the mouse barrel cortex. Individual plots represent the data from a specific hybridisation in which probes from individual animals are co-hybridised with reference sample. The header of each plot is the code assigned to each individual animal, which is described in detail in section 5.3.1. Note how the distributions in each plot approximate a bell-shaped or normal distribution (blue line), centred around $\log_2$ (Experiment/ Reference) ratios = 0, a consequence of applying a log transformation to microarray data.

Similarly the observed data from the microarray hybridisations are plotted and the R-J value for each experiment is calculated. This value is then used to derive a p-value which is used to accept or reject the null hypothesis that the observed data are derived from a normal distribution.

The p-values for each hybridisation tested were lower than the pre-stated, acceptable p-value of 0.01 (table 5.3) and the null hypothesis for each hybridisation is rejected ie none of the $\log_2$ (Exp/Ref) ratios from any hybridisation appear to follow a normal distribution. As the distributions of the $\log_2$ (Exp/Ref) ratios from the hybridisations did not follow a normal distribution, traditional parametric statistical tests, that rely on prior knowledge about the distribution of the data, could not be applied.

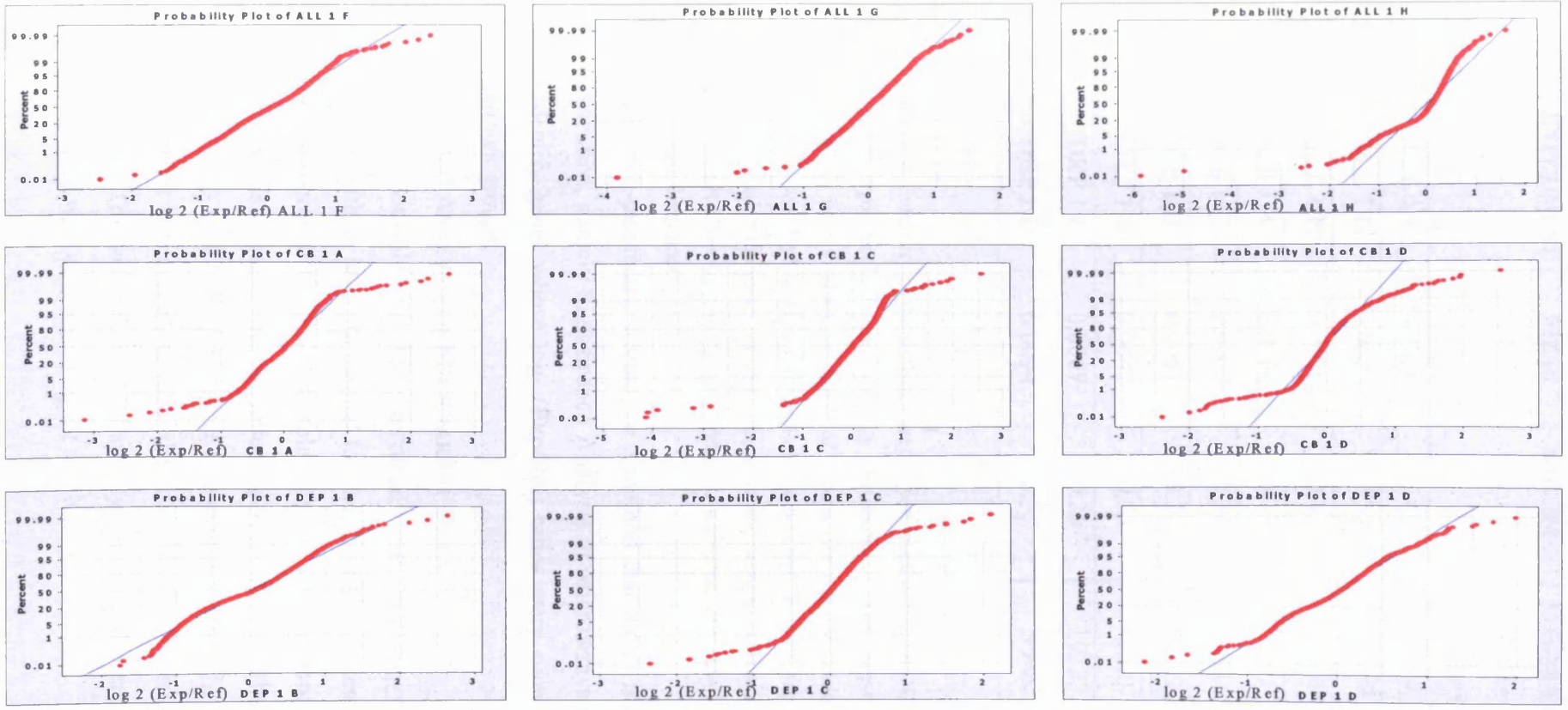


Fig 5.10 Probability plots of data from microarray experiments designed to analyse the effects of vibrissal deprivation, for a period of 1 day, on gene expression in the mouse barrel cortex. Suitably deprived animals were co-hybridised with a reference sample and following data transformation the log2 (Cy5 Exp/ Cy3 Ref) ratios at each spot on the microarray were calculated. The title of each plot represents the sample code used to describe each individual animal (see section 5.3.1 for further details). Plotted are the normal probabilities (y-axis) versus the log2 (Exp/Ref) ratio for the observed data in the bins of the distribution histograms (figure 5.9) for each hybridisation. It is clear that the data depart from the blue line at the extremes of the distribution. The blue line is an indicator of how the data should look if the ratios were drawn from a normally distributed dataset. For all plots the lowest ratios are above the line, while the highest ratios are above the line. This suggests that the distribution of the ratios in these experiments are biased with heavier tails than a normal distribution ie more data points lie at the extremities.

EXPERIMENT	MEAN	STD. DEV.	NUMBER OF DATA POINTS	R-J VALUE	P-VALUE
ALL 1 F	0.0036	0.46	9654	0.997	<0.010
ALL 1 G	0.0091	0.32	10287	0.994	<0.010
ALL 1 H	0.0075	0.42	8952	0.920	<0.010
CB 1 A	0.0064	0.33	9601	0.981	<0.010
CB 1 C	0.0063	0.34	11345	0.971	<0.010
CB 1 D	0.0051	0.27	9783	0.954	<0.010
DEP 1 B	0.0042	0.53	8760	0.992	<0.010
DEP 1 C	0.0067	0.26	10332	0.988	<0.010
DEP 1 D	0.0094	0.37	9946	0.955	<0.010

Table 5.3 A summary of the statistics used to test whether the $\log_2$ (Cy5 Exp/ Cy3 Ref) ratios generated in microarray hybridisations are normally distributed. The experiment name refers to microarray hybridisations used to study the effect of 1 days vibrissal deprivation on gene expression in the mouse barrel cortex with Cy3 labelled reference. For a description of the experiment codes see section 5.3.1. For each hybridisation the mean, standard deviation and number of transformed data points are provided. The results of the R-J test are provided in addition to a p-value. P-values less than 0.01 indicate that the data is not normally distributed.

Rather, non-parametric testing, or a method that didn't make any assumption about the distribution, namely Significance Analysis of Microarrays (SAM, section 3.8), was used to identify genes whose $\log_2$ (Exp/Ref) ratios were significantly different amongst the treatment groups.

5.4.1 SAM ANALYSIS OF MICROARRAY DATA

SAM is available for free download from http://www-stat.stanford.edu/~tibs/SAM/ complete with an in-depth manual for use, tutorials and test datasets. The software works through Microsoft Excel as a simple plug-in. Prior to conducting SAM analysis, the $\log_2$ (Exp/Ref) ratios (available in Appendix 8) for the hybridisations under comparison were collated into an Excel spreadsheet (see figure 5.11a). In order for SAM analysis to be conducted it is necessary to label the columns of the spreadsheet with the corresponding treatment. This was achieved by simply replacing the experiment identifier, in row 1 of figure 5.11a with an integer (figure 5.11b).

a)

NIA 15K C	ALL 1 F	ALL 1 G	ALL 1 H	CB 1 A	CB 1 C	CB 1 D	DEP 1 B	DEP 1 C	DEP 1 D	COUNT
H3008E12	-0.30795	0.20985	0.148005	0.123358	0.088234	0.046043	0.026254	0.209816	0.384488	9
H3016E11	-0.0785	0.208415	0.441204	-0.37686	-0.60386	-0.76579	-0.56996	0.091796	-0.2343	9
H3018E03	0.091789	0.108809	0.056904	-0.03071	-0.16082	0.012198	0.481759	0.44588	0.613648	9
H3018A04	-0.14912	0.197633	0.093382	-0.04505	-0.02113	-0.12487	0.395032	0.092732	0.248701	9
H3028E12	0.597981	0.386955	0.818122	0.383122	0.050074	0.052813	-0.54192	-0.09365	-0.46445	9
H3046A04	0.196122	0.161057	0.472716	0.348775	0.010183	0.017783	-0.38321	-0.03206	-0.214	9
H3051A07	0.849345	0.328611	0.414931	0.505621	0.582629	0.503847	0.57627	0.639163	1.162639	9
H3063A08	-0.47582	-0.24658	0.014173	0.326109	0.491487	0.184802	-0.18188	0.102786	0.58375	9
H3070E04	-0.4363	-0.05288	0.048231	0.450845	0.365435	0.034547	-0.0967	0.124946	0.406436	9
H3087E07	-0.02313	0.146865	0.158614	0.048306	0.084191	0.02096	-0.0861	0.105913	0.06967	9
H3095E07	-0.37494	0.038414	0.110098	0.231779	0.201583	0.094997	-0.16091	0.28182	0.41622	9
H3095E08	-0.2118	-0.13711	-0.17932	0.135686	-0.04827	0.091217	-0.02123	0.07635	0.473357	9
H3104A12	0.365816	0.761829	0.224187	-0.19193	-0.14926	0.083444	0.913949	0.506682	0.654336	9
H3107E08	0.364952	0.37656	0.113305	0.121535	-0.06758	0.300725	0.794806	0.442383	0.679344	9
H3110E03	0.407029	0.463986	0.143682	0.138119	-0.04325	0.287053	0.818148	0.530326	0.722998	9
H3124E12	0.022862	0.070821	0.188243	0.281781	0.027451	0.087128	0.078415	0.178675	0.159114	9
H3126A11	0.711898	0.691891	0.278273	-0.01599	-0.13038	0.445997	1.261043	0.668373	0.889502	9
H3134A03	-0.19433	-0.30752	0.149908	0.50813	0.976835	0.114435	0.198801	0.178727	0.872941	9
H3134A04	0.922999	0.699552	1.18696	0.616595	0.004132	0.148699	-0.5188	0.015839	-0.14766	9
H3140A11	0.765799	0.600296	1.049105	0.50316	-0.07431	0.129016	-0.2414	0.089196	-0.32567	9

COLUMN 1 | COLUMNS 2 3 4 | COLUMNS 5 6 7 | COLUMNS 8 9 10 | COLUMN 11

b)

NIA 15K C	1	1	1	2	2	2	1	1	1
H3008E12	-0.30795	0.20985	0.148005	0.123358	0.088234	0.046043	0.026254	0.209816	0.384488
H3016E11	-0.0785	0.208415	0.441204	-0.37686	-0.60386	-0.76579	-0.56996	0.091796	-0.2343
H3018E03	0.091789	0.108809	0.056904	-0.03071	-0.16082	0.012198	0.481759	0.44588	0.613648
H3018A04	-0.14912	0.197633	0.093382	-0.04505	-0.02113	-0.12487	0.395032	0.092732	0.248701
H3028E12	0.597981	0.386955	0.818122	0.383122	0.050074	0.052813	-0.54192	-0.09365	-0.46445
H3046A04	0.196122	0.161057	0.472716	0.348775	0.010183	0.017783	-0.38321	-0.03206	-0.214
H3051A07	0.849345	0.328611	0.414931	0.505621	0.582629	0.503847	0.57627	0.639163	1.162639
H3063A08	-0.47582	-0.24658	0.014173	0.326109	0.491487	0.184802	-0.18188	0.102786	0.58375
H3070E04	-0.4363	-0.05288	0.048231	0.450845	0.365435	0.034547	-0.0967	0.124946	0.406436
H3087E07	-0.02313	0.146865	0.158614	0.048306	0.084191	0.02096	-0.0861	0.105913	0.06967
H3095E07	-0.37494	0.038414	0.110098	0.231779	0.201583	0.094997	-0.16091	0.28182	0.41622
H3095E08	-0.2118	-0.13711	-0.17932	0.135686	-0.04827	0.091217	-0.02123	0.07635	0.473357
H3104A12	0.365816	0.761829	0.224187	-0.19193	-0.14926	0.083444	0.913949	0.506682	0.654336
H3107E08	0.364952	0.37656	0.113305	0.121535	-0.06758	0.300725	0.794806	0.442383	0.679344
H3110E03	0.407029	0.463986	0.143682	0.138119	-0.04325	0.287053	0.818148	0.530326	0.722998
H3124E12	0.022862	0.070821	0.188243	0.281781	0.027451	0.087128	0.078415	0.178675	0.159114
H3126A11	0.711898	0.691891	0.278273	-0.01599	-0.13038	0.445997	1.261043	0.668373	0.889502
H3134A03	-0.19433	-0.30752	0.149908	0.50813	0.976835	0.114435	0.198801	0.178727	0.872941
H3134A04	0.922999	0.699552	1.18696	0.616595	0.004132	0.148699	-0.5188	0.015839	-0.14766
H3140A11	0.765799	0.600296	1.049105	0.50316	-0.07431	0.129016	-0.2414	0.089196	-0.32567

Figure 5.11 Data formatting of a microarray dataset prior to analysis by SAM software.

a) In column 1 is the NIA 15K microarray gene ID, which is a unique identifier for each spot on the microarray. The $\log_2$ (Expt/Ref) ratios for each spot, in each hybridisation (described in section 5.3.1), are arranged in columns 2-10. A particular hybridisation is identified by the field header. Columns 2, 3 and 4 (red) contain the ratios from the ALL treatment group. Columns 5, 6 and 7 (yellow) contain the ratios from the chessboard (CB) treatment groups while columns 8, 9 and 10 (green) are the all deprived (DEP) treatment group.

Column 11 is a simple COUNT function used to remove genes with missing data in more than 1 hybridisation.

b) The data is now ready for SAM analysis. The count column is removed and the labels for each individual experiment are replaced with an integer that signifies the replicates of all treatments. These are the labels which are swapped around when permutations are applied in the SAM analysis. This is an example of a two-class analysis, in which the treatments ALL and DEP are thought by SAM to be the same (re-labelled 1), while CB treatment is re-labelled 2. The aim of this analysis is to identify genes, with expression levels that are significantly different between CB treatment and both ALL and DEP ie genes whose expression changes only under those conditions that induce neuronal plasticity.

SAM provides two options for analysing the dataset; a multi-class response (for three or more treatments) or a two-class response (two treatments). While it may seem more intuitive to run a multi-class response, as three treatments are being analysed, through trials (data not shown) it was observed that many genes significantly different in two-class response analyses were not being identified in the multi-class analysis. Recalling that the overall aim of this project is to identify genes, whose expression levels are significantly different when the CB treatment is compared to both the ALL and DEP treatments, it was decided to analyse this and other datasets using the two-class response. Thus, as shown in figure 5.11b, the data labels of each treatment group are re-labelled so that SAM deals with, for example, ALL and DEP as a single treatment and these genes are compared to the CB treatment. Therefore three analyses of every dataset were carried out;

ALL and DEP as one treatment vs CB

CB and DEP as one treatment vs ALL

ALL and CB as one treatment vs DEP

Some user input is required in choosing a suitable threshold, Δ, with which to define the number of genes being called significant; a higher Δ will retrieve a smaller list of genes while a lower Δ produces a longer list. Though it may seem better to choose a longer list, the trade-off is an increase in the number of falsely called genes (FDR) as determined by the permutations amongst the labels.

SAM provides two methods for setting the threshold. Both the graphical method (figure 5.12) and the tabular method (table 5.4) can be used interchangeably. In figures 5.12a and b genes called differentially expressed are shown in red and green. The effects of setting a low and high stringency threshold, on the number of genes called significant, are shown in figures 5.12a and b respectively. By manually controlling a slider the threshold can be altered such that more or less points at each extremity are called. This method of setting a threshold may be useful if the graph shows an obvious departure or bending from the middle section, coloured black, which represents genes not significant.

A threshold can also be set by studying the tabular output of SAM (table 5.4). In table 5.4, the number of genes called significant, in ALL and DEP vs CB, are provided, together with the median and 90^{th} percentile FDR. The FDR is expressed as a percentage ie the number of genes falsely called significant (determined by permuting the treatment labels) compared to the number of genes called significant when the treatment labels are

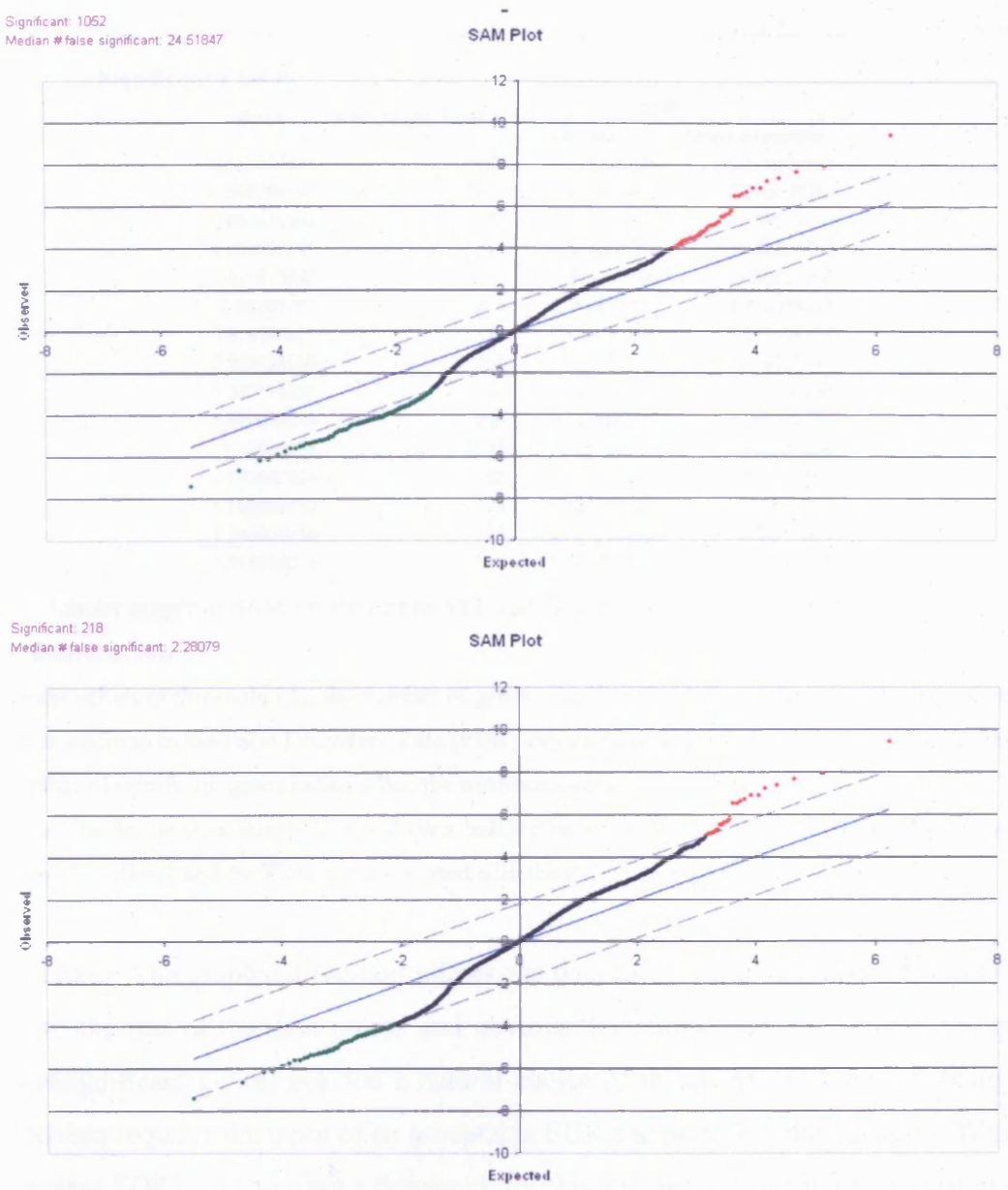


Figure 5.12 Graphical output of SAM for data from ALL and DEP vs CB two-class response analysis for the Cy3 Ref TP1 hybridisations.

For every gene in the set, the d-statistic (similar to a t-statistic) amongst the two treatments is calculated. The process is repeated with the labels switched amongst the treatments. Following each permutation the median (or 90th percentile) d*- values of all permutations are calculated. In the SAM plots a) and b) the expected d*-value (x-axis) are plotted against the actual, observed d-values (x-axis) for each gene. Genes (red and green) above a user defined threshold (delta top right corner, broken lines) are considered significant a) A low stringency threshold is chosen which identifies 1052 significantly differentially expressed genes with a 90th percentile false discovery rate (FDR) of 21% (see also table 5.4). b) A high threshold is set. Notice less coloured spots in the plot compared to a). This value of delta has found 218 significant genes with a FDR of 2.1% genes (see also table 5.4).

Significance Table

Delta	# of Genes Called	FDR (%)	
		Median	90th percentile
0.553099641	1282	1.776624679	22.79216817
0.598284199	1041	1.212586491	20.46107901
0.658475208	838	0.81865631	15.98017117
0.698391741	717	0.574087019	13.57907162
0.821813637	560	0.245012138	5.958695206
0.85207761	516	0.212723717	4.754375077
0.915386851	427	0.128530958	3.637426107
0.982825165	357	0.153733106	3.09003544
1.040714238	282	0.194619571	2.374358765
1.087669838	220	0.249466905	2.245202141
1.10895828	193	0.28436642	1.990564938
1.152947824	153	0.358710582	1.829423967
1.199356499	104	0.527718452	2.110873808
1.249824636	73	0.751818069	1.578817944
1.314898018	45	1.219615978	2.439231956

Table 5.4 Tabular output of SAM for data from ALL and DEP vs CB two-class response analysis for Cy3 Ref TP1 hybridisations.

For different values of threshold (Δ), the number of genes, significantly different amongst the treatments, is provided, in addition to the False Discovery Rate (FDR) expressed as a number of genes or as a percentage of the number of significant genes called when the treatments are not permuted.

The user has to decide on a threshold based on a balance between the number of differentially expressed genes called significant and the FDR rate associated with the chosen threshold (see also figure 5.12).

not permuted. The graphical method allows the data to be visualised and interpreted in relation to the rest of the data points and obvious deviations from the middle group of points (insignificant genes) provide a natural choice of threshold. In contrast using the tabulated data requires the input of an acceptable FDR and provokes the question; What is an acceptable FDR? In choosing a threshold, for this and every other analysis carried out, it was decided that rather than using the tabulated data, the graphical method described previously would be used.

All hybridisations in which the reference sample was labelled with Cy3 worked well. However the Cy5 Reference series of experiments suffered from a number of failed hybridisations. A total of nine hybridisations failed to generate data worthy of analysis (figure 5.13) principally due to very high background contamination the exact source of which was unknown. As the contamination was different in appearance to dust particles and the microarrays and probes were generated and hybridised in parallel with good quality arrays it is possible that the slides were contaminated during manufacture by the supplier. Experimental failure had downstream effects on data analysis as valuable biological replication was lost and certain two class comparisons could not be carried out

eg for Cy5 Ref time point 8 no undeprived controls were available, therefore the only available comparison was between CB and DEP (see table 5.5).

Cy3 Ref	Cy5 Ref
ALL 0.4 A	ALL 0.4 A
ALL 0.4 B	ALL 0.4 B
CB 0.4 A	CB 0.4 A
CB 0.4 C	CB 0.4 C
CB 0.4 D	CB 0.4 D
DEP 0.4 A	DEP 0.4 A
DEP 0.4 B	DEP 0.4 B
DEP 0.4 C	DEP 0.4 C
ALL 1 F	ALL 1 F
ALL 1 G	ALL 1 G
ALL 1 H	ALL 1 H
CB 1 A	CB 1 A
CB 1 C	CB 1 C
CB 1 D	CB 1 D
DEP 1 B	DEP 1 B
DEP 1 C	DEP 1 C
DEP 1 D	DEP 1 D
ALL 2 F	ALL 2 F
ALL 2 G	ALL 2 G
ALL 2 H	ALL 2 H
CB 2 A	
CB 2 D	
CB 2 F	CB 2 F
DEP 2 B	
DEP 2 D	DEP 2 D
DEP 2 F	DEP 2 F
ALL 4 F	ALL 4 F
ALL 4 G	ALL 4 G
ALL 4 H	
CB 4 C	CB 4 C
CB 4 E	CB 4 E
CB 4 F	
DEP 4 F	DEP 4 F
DEP 4 G	DEP 4 G
DEP 4 I	DEP 4 I
ALL 8 A	
ALL 8 B	
ALL 8 C	
CB 8 A	CB 8 A
CB 8 E	CB 8 E
CB 8 H	CB 8 H
DEP 8 B	DEP 8 B
DEP 8 D	DEP 8 D
DEP 8 E	DEP 8 E

Table 5.13 Hybridisations from which data worthy of analysis could be extracted

The results of SAM analysis are provided in table 5.5. For all two-class comparisons carried out the number of transcripts analysed are provided in addition to the number of significantly differentially expressed transcripts observed. The FDR determined by SAM is presented and was below 10.5% for all analyses that identified differentially expressed transcripts. Technical replications of each experiment were analysed separately such that three lists of significant transcripts were generated for Cy3 Ref and Cy5 Ref at time points 0.4, 1, 2, 4 and 8 days respectively (in reality some comparisons did not identify any differentially expressed transcripts, therefore the respective list was empty). The three lists were collated and, in cases where the same transcript was significant in more than one comparison, replicates were removed. Using this method the numbers of transcripts differentially expressed at each time point in separate technical replications was calculated (table 5.5). The lists of significant transcripts for each technical replicate were then compared and those that replicated were identified, the numbers are shown in table 5.5.

For the 1 day deprivation experiments large numbers of genes were called significant in each replication. In Cy3 Ref TP 1 experiments 68% of genes called significant were common to both replications, while only 41% of genes from the Cy5 Ref experiments were in the intersection of the two experiments. One of the strengths of SAM is that the FDR can be controlled. Calling fewer genes significant results in a lower FDR, therefore it was investigated whether reducing the number of genes, called significant, in Cy3 Ref TP 1 and Cy5 Ref TP 1 experiments, would affect the percentage of genes common to both lists. Reducing the number of genes called significant to approximately 100 reduced the average FDR to 0.1% for each replicated experiment. This made no significant difference to the percentage of genes that were common however. This does lend weight to the argument that microarray experiments can generate many false positives and adequate replication is essential to remove random associations.

In order to visualise the differences in expression amongst the treatment groups, the genes identified as significant for all time points will be investigated in section 5.4.2 using cluster analysis.

Technical Relication	Length of Deprivation (days)	No. of Transcripts in Analysis	Analysis	No. of Transcripts differentially expressed in each analysis	Median False Discovery Rate (%)	No. of Transcripts differentially expressed when analyses are combined and filtered for replicates	No. of differentially expressed transcripts common in replicated experiments
Cy3 Ref	0.4	7118	CB	0	NA	395	0
			DEP	0	NA		
			ALL	395	6.5		
Cy5 Ref		6979	CB	0	NA	0	
			DEP	0	NA		
			ALL	0	NA		
Cy3 Ref	1	9599	CB	256	4.4	528	361
			DEP	484	1.6		
			ALL	299	7.4		
Cy5 Ref		6747	CB	163	8	879	
			DEP	778	0.4		
			ALL	511	0.3		
Cy3 Ref	2	10683	CB	0	NA	0	0
			DEP	0	NA		
			ALL	0	NA		
Cy5 Ref		0	CB	F	NA	0	
			DEP	F	NA		
			ALL	F	NA		
Cy3 Ref	4	10191	CB	158	2.8	584	164
			DEP	0	NA		
			ALL	536	9.7		
Cy5 Ref		5313	CB	476	1.3	603	
			DEP	0	NA		
			ALL	301	10.3		
Cy3 Ref	8	8651	CB	0	NA	618	92
			DEP	166	8.1		
			ALL	554	0.8		
Cy5 Ref		4324	CBvsDEP	267	7.7	267	

Table 5.5 Results of a SAM analysis of microarray data. Transcripts significantly differentially expressed in the mouse barrel cortex following vibrissal deprivation for defined periods of time (column 2) were identified using SAM analysis. At each time point the number of transcripts above background on the microarrays and analysed in three two-class comparisons are provided. F – Experimental fails Analysis codes are Analysis CB = CB vs DEP and ALL Analysis DEP = DEP vs ALL and CB Analysis ALL = ALL vs CB and DEP, for Cy5 Ref 2 and 8 days deprivation experimental failure meant that only ALL vs DEP and CB vs DEP comparisons respectively could be carried out. The median False Discovery Rate (number of falsely called differentially expressed transcripts expressed as a percentage of the number of true differentially expressed transcripts) is given for each analysis. Following the three SAM analyses significant transcript lists were collated and filtered for replicates, the numbers are provided in column 7. Finally the collated transcript lists were compared when each experiment was technically replicated. The number of significant transcripts that could be replicated for each time point are shown in column 8.

5.4.2 CLUSTER ANALYSIS OF MICROARRAY DATA

The sheer volume of data generated in microarray experiments means that the analysis of each gene individually is counter-productive. However, a list of genes showing differential expression in a particular experiment, answers very few questions about the scientific problem being investigated. To this end, investigators using microarray technology use a variety of methods to identify patterns in the data, providing an overview of it.

Typical approaches include the data reduction methods Principal Component Analysis, Multi-Dimensional Scaling and cluster analysis (Slonim, 2002). Cluster analysis, in particular, has been used in the majority of microarray publications and is an unsupervised method for identifying patterns in complex datasets. No assumptions are made regarding the distributions of the data and clustering, in contrast to statistical testing, identifies data with similar properties (Slonim, 2002) therefore, it should be noted that clustering assigns no level of confidence to the results generated. Clustering can be carried out on any dataset eg genes, treatments, timepoints etc. and it is this flexibility that has seen it being used in various ways in microarray experiments, including defining disease sub-classes (Golub et al., 1999) and attempting to identify the function of uncharacterised genes (Eisen et al., 1998).

In hierarchal clustering, each group of data start as individual points and the two data groups, that are most closely related, are merged to form a cluster. This relation is determined by a similarity metric which can be any property of the data eg the expression levels of two genes measured in n samples. The process is repeated to find the most related clusters, to form new clusters of 4 samples. The process is repeated until a single cluster remains. Thus the data is arranged into a tree with various branches representing clusters (eg figure 5.15). The tree can be cut at any point resulting in more, or less clusters. By cutting the tree so that more clusters are identified, more similar groups of data are identified (Slonim, 2002).

Here, hierarchal clustering will be used to visualise patterns, in the expression profiles of genes identified as significantly differentially expressed, in the mouse barrel cortex, when various treatments were applied to the vibrissae over a period of 8 days.

Hierarchal clustering was carried out on the lists of genes, identified as significantly differentially expressed, in both technical replicates (where possible), in the mouse barrel cortex following vibrissal deprivation. Imagene™ (Biodiscovery, Inc. CA, USA) microarray analysis software was used to carry out the clustering and the results, for each

time point, are presented in figures 5.15 – 5.19. In each cluster diagram, the data from individual hybridisations are arranged in rows. The rows are separated into three groups representing the treatments applied to the vibrissae. Within each treatment group, the technical replication experiment data is shown in separate rows, Cy3 Ref followed by Cy5 Ref. A particular hybridisation can be identified by the label at the bottom of each column.

The colours in the cluster diagram represent the relative expression levels (log_2 Experiment/Reference) of each gene, where red indicates high values and green low values (scale is given, bottom left).

Effects of 0.4 days Vibrissal Deprivation

The hybridisations conducted to analyse the effects of 0.4 days vibrissal deprivation, all generated data worthy of analysis (figure 5.13). This experiment was successfully replicated in a technical dye swap. However, in a SAM analysis of the Cy5 Ref data no genes were identified as significantly differentially expressed.

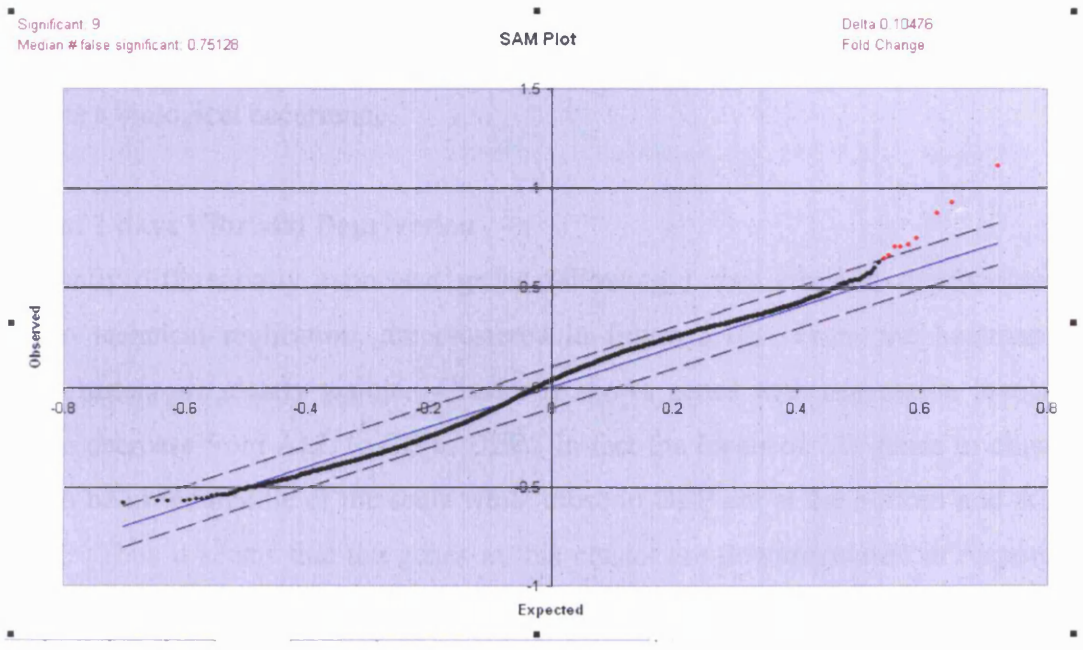


Figure 5.14 Example of the SAM plot when no differential gene expression is observed.
For every gene compared the actual d statistic (y-axis) is plotted against the median permuted d* value. When no differential expression is observed d = d* and the line shows no departure from this relationship at the extreme tails. This figure can be compared with figure 5.16, which is a typical SAM plot when differentially expressed genes are found.

Figure 5.14 is an example of the graphical output from the SAM analysis of the 0.4 day deprivation Cy5 Ref experiments. The graph reveals the reason that SAM failed to identify any significant genes. Recalling that SAM calculates a d-test statistic (similar to

a *t*- statistic) for each gene then does the same when the treatment labels are switched amongst the samples. The treatment labels are switched, or permuted, a number of times and the median d* value is calculated for each gene. Thus genes that are truly differentially expressed will be robust to the permutations and d* will not be equal to d.
In figure 5.14 the d* values are plotted against the d values for each gene. As can be seen all the data points show correlation and a straight line results. Thus for each gene in the Cy5 Ref data set d* is equal to d and no genes are being called as differentially expressed. This figure can be compared with figure 5.12 in which differentially expressed genes are observed and the graph departs from y = x, or d* = d, in the extreme tails, allowing significant genes to be identified.
The cluster analysis appears to show a single cluster of genes with a profile that suggests that vibrissal deprivation is down regulating their expression. The expression profiles of the genes declared significantly differentially expressed following 0.4 days deprivation in the Cy3 Ref experiments are presented in figure 5.15.
The failing of the Cy5 Ref experiments to generate any significant genes makes it highly unlikely that the expression profile of the cluster observed in figure 5.15 accurately represents a biological occurrence.

Effects of 1 days Vibrissal Deprivation

Significantly differentially expressed genes following 1 day vibrissal deprivation and robust to technical replication, are clustered in figure 5.16. From the heatmap two distinct clusters are clearly visible. Cluster A shows genes with expression levels that appear to decrease from ALL to CB to DEP. In fact the levels of CB genes in cluster A appear to be in the middle of the scale while those in DEP are at the bottom and ALL at the TOP. Thus it seems that the genes in this cluster are downregulated in response to vibrissal deprivation. Furthermore, the degree of downregulation appears to increase as more vibrissae are removed (DEP removes all vibrissae, CB removes half).
The genes in cluster B show a markedly different profile (see also figure 5.17). Here it appears that 103 genes are upregulated following CB treatment in comparison to both ALL and DEP treatments. It therefore appears that differential expression of these genes is associated only with conditions that induce neuronal plasticity. A thorough examination of the gene list is provided in chapter 6. A comprehensive list of the genes in clusters A and B are provided in Appendix 10 under the file names Cy3Cy5 Ref time

point 1 Activity and Cy3Cy5 Ref time point 1 Chessboard respectively and for cluster B in table 5.6.

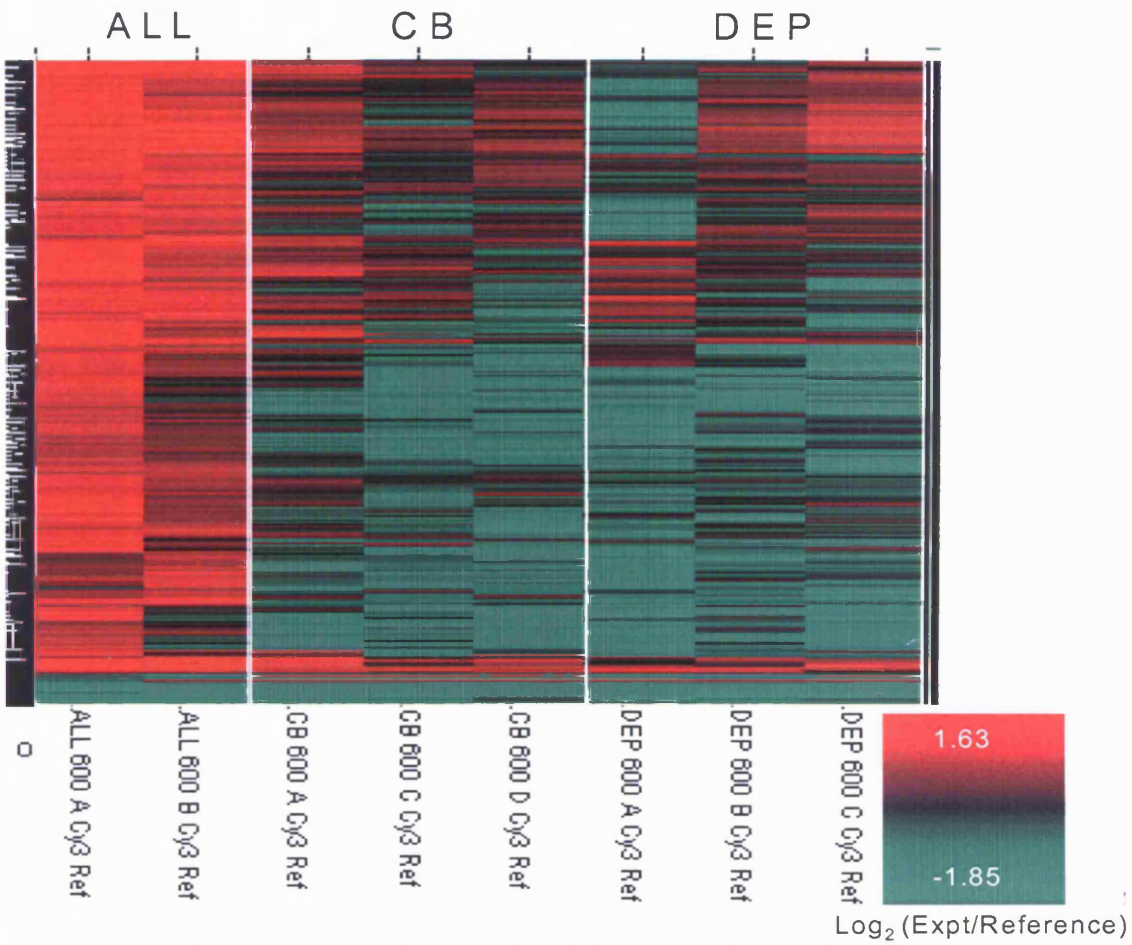


Figure 5.15 The Hierarchal Clustering of Genes Differentially Expressed in the Mouse Barrel Cortex following 0.4 Day Vibrissal Deprivation

Cluster analysis is used to identify genes with similar expression profiles in the mouse barrel cortex when the vibrissae are subjected to one of three treatments (ALL, CB or DEP) for 0.4 day. In the heat map the relative expression levels of each gene are plotted in rows. The numerical values are assigned a colour, with red signifying high values and green low (see scale, bottom left). Genes with similar expression in all samples are clustered together. The data is arranged in columns, with each column representing a particular hybridisation (label at bottom of heatmap). The columns are grouped according to the treatment applied. Each group is made up of columns representing the data from the biological replicates. The general effect seen here is of higher expression in the ALL treatment group compared to both CB and DEP.

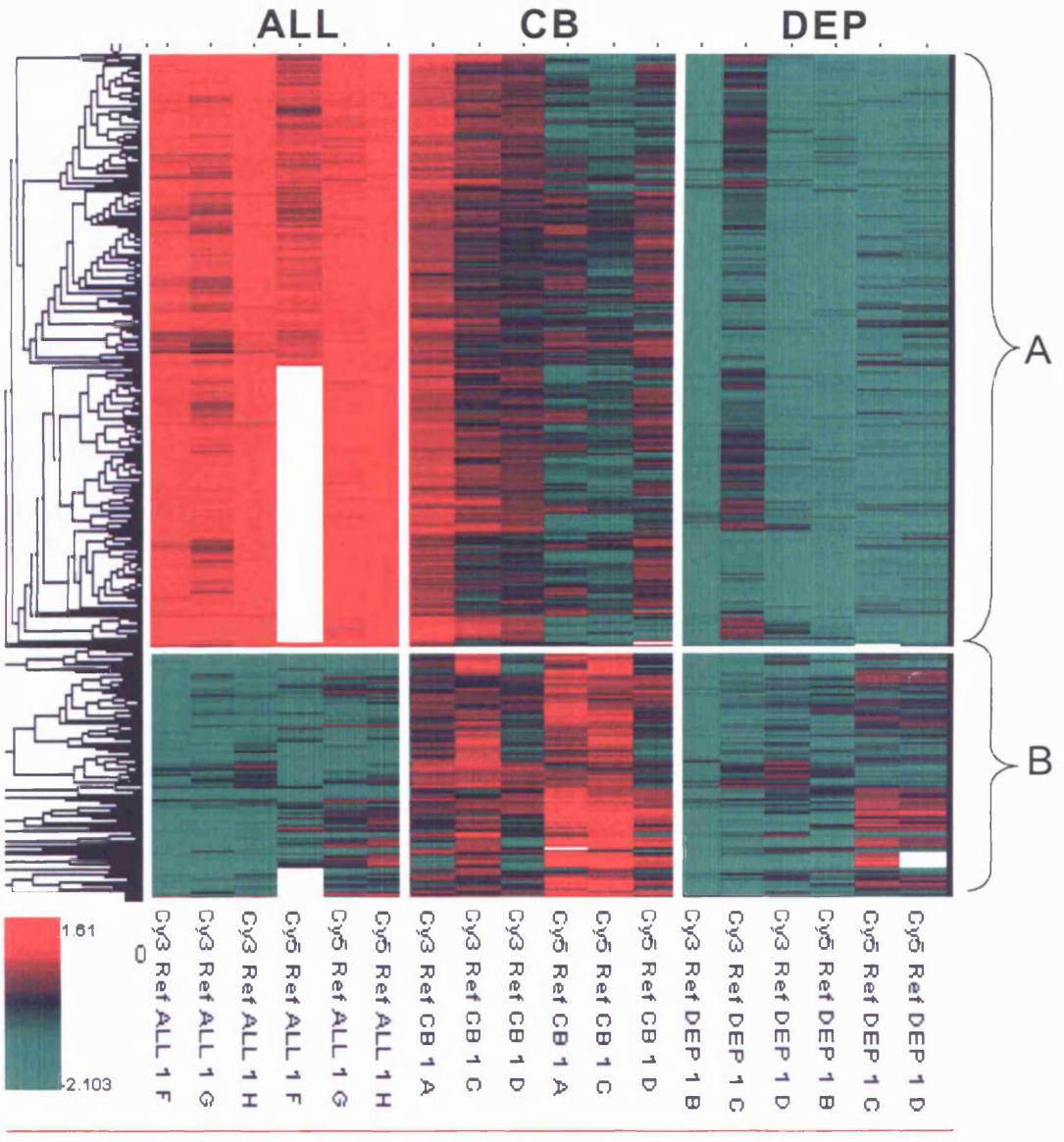


Figure 5.16 The Hierarchal Clustering of Genes Differentially Expressed in the Mouse Barrel Cortex following 1 Day Vibrissal Deprivation

Cluster analysis is used to identify genes with similar expression profiles in the mouse barrel cortex when the vibrissae are subjected to one of three treatments (ALL, CB or DEP) for one day. In the heat map the relative expression levels of each gene are plotted in rows. The data is arranged in columns, with each column representing a particular hybridisation (label at bottom of heatmap). The columns are grouped according to the treatment applied (label top of diagram). Each group is made up of six columns representing hybridisations form three biological replicate hybridisations and a technical replication or dye swap experiment. Cluster A - Genes appear to be downregulated with both deprivation treatments. Cluster B- Genes appear to be upregulated following CB deprivation when compared to both ALL and DEP treatment.

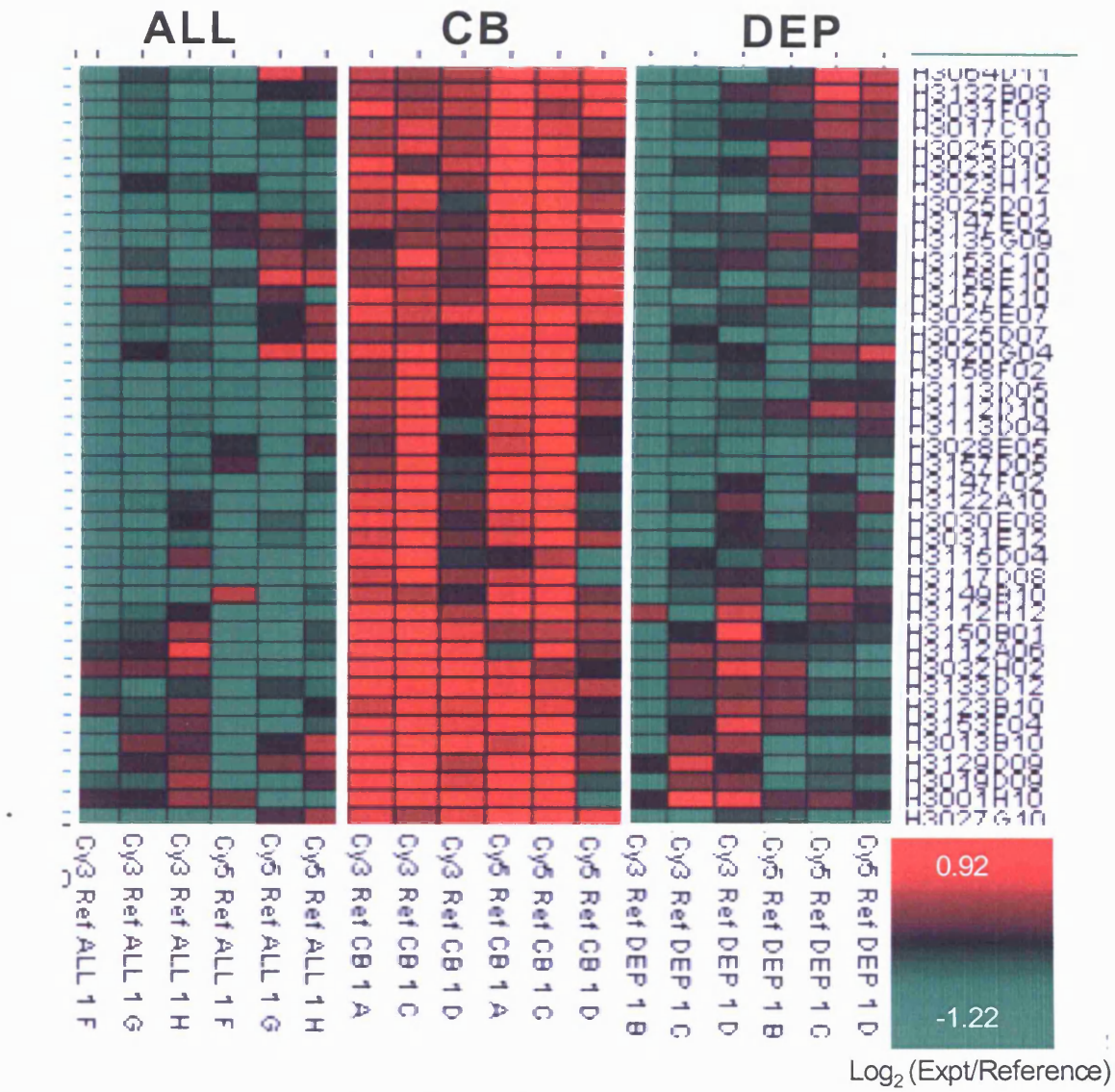


Figure 5.17 Genes upregulated following 1 day chessboard deprivation.

Enlarged version of of the heatmap shown in figure 5.16. This figure reveals the expression profiles of individual genes in a cluster that appear upregulated in the barrel cortex following chessboard deprivation. The samples are arranged in rows as described in figure 5.16. High expression levels are coloured red while low expression levels are green, for scale see bottom right.

Effects of 2 days Vibrissal Deprivation

In these experiments a high failure rate prevented any statistical analysis from being carried out on the data from the Cy5 ref experiments. This does not however have any bearing on results as no genes were identified as significantly differentially expressed in the Cy3 Ref experiments. Again, the SAM plot showed a flat line (as in figure 5.14) and therefore no genes were called.

NIA 15K Clone Set ID	Putative Gene ID	Gene name	Fold Change CB ALL	Fold Change CB DEP
H3146G06	aconitase 2, mitochondrial (Aco2), mRNA	Aco2	1.622	1.431
H3139G09	similar to protein-tyrosine-phosphatase [] (LOC264656), mRNA	Acp1	1.417	1.374
H3025E07	ADP-ribosylarginine hydrolase	Adprh	1.328	1.421
H3112H12	Similar to , clone IMAGE:1349076, mRNA	Adrbk1	1.517	1.289
H3147E02	S-adenosylmethionine decarboxylase 1 (Amd1), mRNA	Amd1	1.317	1.260
H3114D08	actin related protein 2/3 complex, subunit 3 (21 kDa)	Arpc3	1.564	1.267
H3147F02	ash2 (absent, small, or homeotic)-like (Drosophila) (Ash2l), mRNA	Ash2l	1.555	1.350
H3020G04	basigin (Bsg), mRNA	Bsg	1.213	1.289
H3125D07	Similar to LOC241248 [LOC241248], mRNA	Bup	1.484	1.441
H3031D11	similar to Calcineurin-binding protein Cabin 1 (Calcineurin inhibitor) (CAIN) (LOC216110), mRNA	Cabin1	1.462	1.349
H3017C10	UNKNOWN	Car	1.326	1.252
H3030F07	coatomer protein complex, subunit beta 2 (beta prime) (Copb2), mRNA	Copb2	1.668	1.379
H3032C01	cytochrome c, somatic	Cycs	1.582	1.749
H3006B05	dystroglycan 1 (Dag1), mRNA	Dag1	1.427	1.116
H3027F02	similar to ASPARTYL-TRNA SYNTHETASE (ASPARTATE--TRNA LIGASE) (ASPRS) (LOC226414), mRNA	Dars	1.417	1.391
H3031F11	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 15 (RNA helicase A) (Ddx15), mRNA	Ddx15	1.510	1.438
H3031C10	Similar to FK506 binding protein 8 (38 kDa), clone MGC:5831 IMAGE:3601511, mRNA, complete cds	Fkbp8	1.502	1.324
H3023H10	ferritin light chain 1 (Ftl1), mRNA	Ftl1	1.497	1.256
H3023H11	ferritin light chain 1 (Ftl1), mRNA	Ftl1	1.600	1.279
H3032D08	Gcgap-pending	Gcgap	1.206	1.217
H3027D10	guanine nucleotide binding protein, beta 2, related sequence 1 (Gnb2-rs1), mRNA	Gnb2-rs1	1.515	1.363
H3009D10	H2A histone family, member Z (H2afz), mRNA	H2afz	1.672	1.803
H3009F03	H2A histone family, member Z (H2afz), mRNA	H2afz	1.715	1.886
H3027E02	H2A histone family, member Z (H2afz), mRNA	H2afz	1.586	1.697
H3019D09	histidine triad nucleotide binding protein (Hint), mRNA	Hint	1.315	1.324
H3023H12	lactate dehydrogenase 1, A chain (Ldh1), mRNA	Ldh1	1.418	1.380
H3135G09	mitogen activated protein kinase kinase 2, full insert sequence	Map2k2	1.260	1.193
H3025D07	mRNA for Ki-67	Mki67	1.260	1.256
H3043D12	Similar to NCK-associated protein 1	Nckap1	1.310	1.309
H3086F10	NADH dehydrogenase (ubiquinone) 1 beta subcomplex 5 (Ndufb5), mRNA	Ndufb5	1.486	1.416
H3133D04		Ndufv2	1.703	1.444
H3010F07	programmed cell death 6 interacting protein	Pdcd6ip	1.721	1.864
H3158F02	prefoldin 5 (Pfdn5), mRNA	Pfdn5	1.371	1.297
H3019C10	Similar to Rattus norvegicus Prohibitin (Phb), mRNA	Phb	1.334	1.257
H3126B08	phospholipase A2, group VII (platelet-activating factor acetylhydrolase, plasma)	Pla2g7	1.716	1.171
H3021G04	protein phosphatase 1, catalytic subunit, alpha isoform (Ppp1ca), mRNA	Ppp1ca	1.411	1.448
H3149B10	protein kinase C, zeta (Prkcz), mRNA	Prkcz	1.418	1.173
H3124F03	prothymosin alpha (Ptma), mRNA	Ptma	1.734	1.376
H3124A02	pleiotrophin (Ptn), mRNA	Ptn	1.971	1.282
H3025D01	reduced expression 3 (Rex3), mRNA	Rex3	1.387	1.392
H3113D04	ribosomal protein L12 (Rpl12), mRNA	Rpl12	1.415	1.248
H3113D05	ribosomal protein L12 (Rpl12), mRNA	Rpl12	1.417	1.338
H3030E08	ribosomal protein L26 (Rpl26), mRNA	Rpl26	1.400	1.393
H3126H12	ribosomal protein L26 (Rpl26), mRNA	Rpl26	1.769	1.483
H3112D10	ribosomal protein, large P2 (Rplp2), mRNA	Rplp2	1.422	1.250
H3150B01	ribosomal protein S15 (Rps15), mRNA	Rps15	1.367	1.230
H3009F01	ribosome binding protein 1 (Rrbp1), mRNA	Rrbp1	1.599	1.659
H3117H11	Similar to retinoid X receptor beta	Rxrb	1.232	1.171
H3112A06	SAR1a gene homolog (S. cerevisiae) (Sara), mRNA	Sara	1.339	1.266

NIA 15K Clone Set ID	Putative Gene ID	Gene name	Fold Change CB ALL	Fold Change CB DEP
H3152F04	selenoprotein P, plasma, 1 (Sepp1), mRNA	Sepp1	1.496	1.521
H3032H02	serologically defined colon cancer antigen 28 (Sdccag28), mRNA	Stard10	1.379	1.301
H3033G01	Scgn10 like-protein (Sclip), mRNA	Stmn3	1.618	1.289
H3142G05	telomerase associated protein 1 (Tep1), mRNA	Tep1	1.833	1.382
H3057G12	transforming growth factor beta 1 induced transcript 4 (Tgfb1i4), mRNA	Tgfb1i4	1.524	1.318
H3066H08	transforming growth factor beta 1 induced transcript 4 (Tgfb1i4), mRNA	Tgfb1i4	1.481	1.288
H3097C06	transforming growth factor beta 1 induced transcript 4 (Tgfb1i4), mRNA	Tgfb1i4	1.417	1.543
H3001H10	thymosin, beta 10 (Tmsb10), mRNA	Tmsb10	1.369	1.182
H3026G06	tumor rejection antigen gp96 (Tra1), mRNA	Tra1	1.773	1.392
H3005B11	UNKNOWN	Txnl	1.741	1.441
H3109B06	ubiquitin-like 3 (Ubl3), mRNA	Ubl3	1.499	1.710
H3013B10	UNKNOWN	Ubxdc2	1.293	1.332
H3129G09	ubiquitin carboxy-terminal hydrolase L1 (Uchl1), mRNA	Uchl1	1.419	1.273
H3031F01	ubiquinol-cytochrome c reductase core protein 1 (Uqcrc1), mRNA	Uqcrc1	1.485	1.320
H3133D12	valosin containing protein (Vcp), mRNA	Vcp	1.384	1.250
H3027G09	similar to zinc finger protein 22 (KOX 15)	Zfp422	1.733	1.311
H3027G10	zinc finger protein 94 (Zfp94), mRNA	Zfp94	1.524	1.510
H3136D02	LOC212371 (LOC212371), mRNA	9430041O17Rik	1.228	1.196
H3158E10	similar to F-box only protein 21, isoform 1, F-box protein 21 (LOC231670), mRNA		1.271	1.363
H3157D10	similar to KIAA1787 protein [Homo sapiens] (LOC216860), mRNA	0610025P10Rik	1.260	1.264
H3137B09	similar to Rho guanine nucleotide exchange factor 4 isoform a	9330140K16Rik	1.652	1.452
H3153F04	similar to RIKEN cDNA 2210412D01 (LOC216169), mRNA		1.418	1.228
H3157D05	Similar to , similar to Hypothetical protein KIAA0036 (LOC209555), mRNA	6820449I09Rik	1.379	1.474
H3151E09	Similar to , clone IMAGE:4036503, mRNA	2010109N14Rik	1.743	1.321
H3148G02	Mus musculus, hypothetical protein, clone MGC:36488 IMAGE:5363354, mRNA, complete cds		1.564	1.380
H3114H06	UNKNOWN	C920011G20Rik	1.459	1.407
H3060G09	UNKNOWN	1810015H18Rik	1.136	1.461
H3123B10	UNKNOWN	2410004F01Rik	1.254	1.213
H3027G02	UNKNOWN	3110057M17Rik	1.517	1.416
H3030D04	UNKNOWN	A730090N16Rik	1.666	1.415
H3045E08	UNKNOWN	BG066574	1.259	1.458
H3067H12	UNKNOWN	BG068658	1.434	1.324
H3123F06	UNKNOWN	BG073525	1.441	1.316
H3122A10		0610006O17Rik	1.500	1.374
H3031E12		0710001O03Rik	1.446	1.320
H3009G09		1110008P14Rik	1.606	1.653
H3028E05		1500002I11Rik	1.358	1.379
H3129D09		1700034P14Rik	1.263	1.304
H3117D08		2210402A09Rik	1.635	1.291
H3126B04		2210402A09Rik	1.638	1.292
H3092B04		2310024K08Rik	1.465	1.496
H3058B08		2310037I24Rik	1.345	1.223
H3153C10		2410091N08Rik	1.292	1.234
H3004F08		2510019J09Rik	1.916	1.469
H3056B04		2510019J09Rik	1.919	1.424
H3025D03		2610002J02Rik	1.435	1.339
H3064D11		BG068342	1.319	1.264
H3004G04		BG075853	1.875	1.363
H3050B10		BG076338	1.665	1.408
H3132B08		D18Wsu98e	1.382	1.185

Table 5.6This table lists 103 transcripts identified by SAM analysis as significantly unregulated in 1 day chessboard deprived compared to all deprived and undeprived barrel cortices. For each transcript a unique NIA 15K clone set ID is provided in addition to putative gene IDs and names. The degree of upregulation was determined by dividing the averaged relative expression levels of each transcript in CB cortices by expression levels in ALL deprived and unDEPrived cortices.

Effects of 4 days Vibrissal Deprivation

The results of microarray experiments are presented in figure 5.18. As can be seen from table 5.5 approximately 600 genes were called significant in both replications of the dye swap experiment. A list of 164 genes were common when the individual experiments were combined, representing approximately 28% of the numbers obtained without technical replication. This is reflective of the data from the time point 1 day experiments in which large numbers of genes are being called significant, only to be excluded when the data from technical replication are compared. Though it was not tested it is unlikely that increasing the stringency of SAM, that is choose fewer genes with a lower FDR, will reduce the percentage of these false positives in the data (section 5.4.1).

The expression profiles of the Cy3 Ref and Cy5 Ref common genes (figure 5.18) are dominated by the high expression levels of the genes (red colour) in the ALL treatment group compared to the lower expression levels of both the CB and DEP (black – red colour). This suggests that the genes in this cluster are being down regulated in response to vibrissal deprivation, though the gradual decrease from ALL to CB with a further decrease in DEP, observed for timepoint 1 is not apparent (figure 5.16)

Of note is a small cluster, marked with an arrow, at the bottom of figure 5.18. These genes may represent genes upregulated with CB treatment compared to ALL and DEP. It should be noted that not all hybridisations provided data for the technical replication and though it may be prudent to carry out the extra hybridisations in order to generate more data on this cluster time constraints prevent this.

Finally one experiment appears as an obvious anomaly in these experiments. The experiment Cy5 Ref Dep 4 (3^{rd} column from the right, figure 5.18) shows an almost constant expression level for each gene despite a normal RI plot and distribution (observed). This does show that with only few replicates anomalous data does become quite apparent.

Effects of 8 days Vibrissal Deprivation

A complete statistical analysis of the Cy5 ref experiments for 8 days vibrissal deprivation could not be carried out as the ALL treatment group hybridisations were of too low quality to extract suitable data. However, a comparison of the relative expression levels following DEP and CB treatments could be carried out.

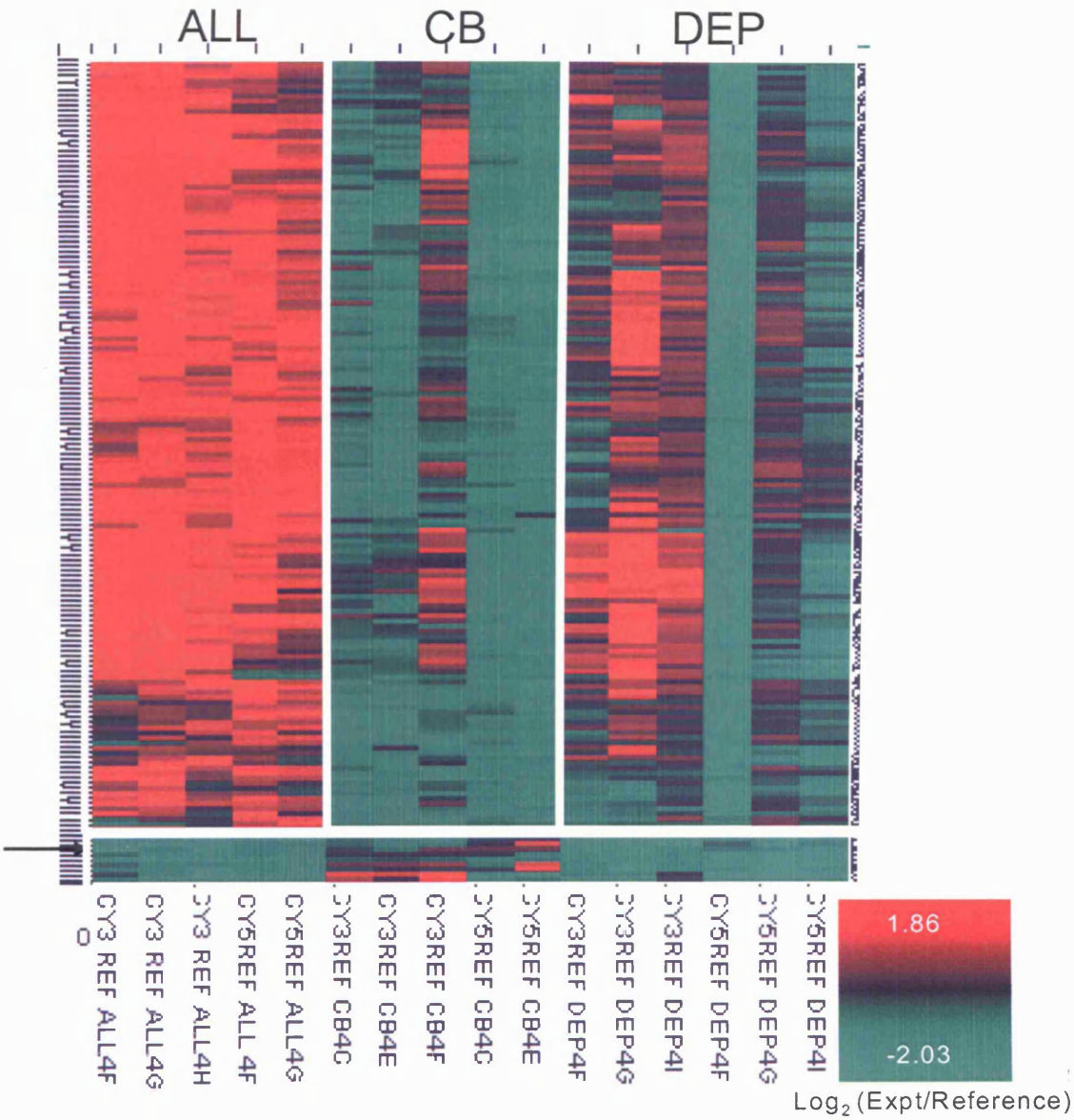


Figure 5.18 Cluster Analysis of Genes Significantly Different following 4 days vibrissal deprivation.
The expression profiles of the genes declared significantly differentially expressed amongst the treatment groups ALL, CB and DEP when the vibrissae were deprived for 4 days. The data is for each gene in a common list ie genes called significantly differentially expressed amongst the samples in both replications of a dye swap.

Of the 267 genes identified, 92 were also declared significant in the Cy3 Ref hybridisations (table 5.5). The relative expression profiles of these genes are presented in figure 5.19. Here, a tendency for the genes to be downregulated when all the vibrissae are deprived (DEP) compared to CB and ALL treatments is observed. Furthermore, the degree of downregulation in CB appears to be approximately half of that in DEP. Though repeating the failed hybridisations to generate data for the ALL treatment group may provide further evidence for this observed effect it is highly unlikely that the extra data

will identify any genes whose expression is regulated exclusively under CB treatment as no such genes were observed in the Cy3 Ref data.

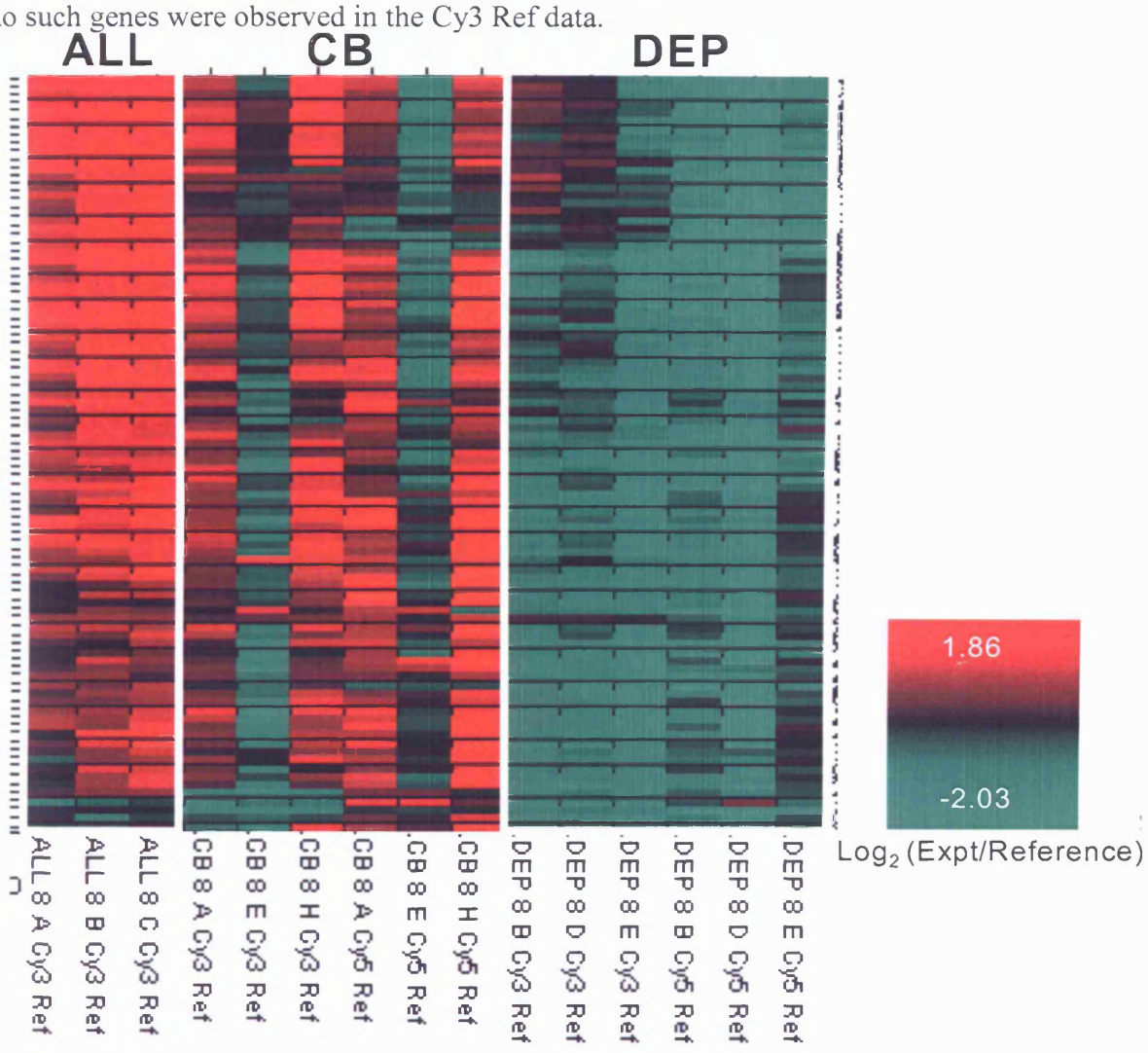


Figure 5.19 Cluster Analysis of Genes Significantly Different following 8 days vibrissal deprivation.
The expression profiles of the genes declared significantly differentially expressed amongst the treatment groups ALL, CB and DEP when the vibrissae were deprived for 8 days. The data is for each gene in a common list ie genes called significantly differentially expressed amongst the samples in both replications of a dye swap. However due to a lack of data for the ALL treatment in Cy5 Ref experiments only those genes significant between the treatments CB and DEP could be identified.

5.5 qPCR VERIFICATION OF MICROARRAY DATA

qPCR (quantitative PCR) is a modification of the PCR reaction (Mullis and Faloona, 1987) used to determine the copy number of specific transcripts in cDNA samples. Fluorescent probes, in this case TaqMan® probes, are added to the reaction mix in an assay that is based on the detection and quantification of fluorescence generated during the reaction (Lee at al., 1987, Livak et al., 1995). TaqMan® probes are designed to be

complementary to sequences within the anti-sense strand of the desired fragment. A reporter dye (FAM) is attached to the 5' end of the probe and a quenching dye (TAMRA) is attached to the 3' base (see figure 5.20). When irradiated the reporter transfers it's energy to the quenching dye via FRET (Forster Resonance Energy Transfer, Hiyoshi and Hosoi, 1994, Chen at al., 1997). Therefore, due to the close proximity of the dyes in an intact probe, there is no fluorescence.

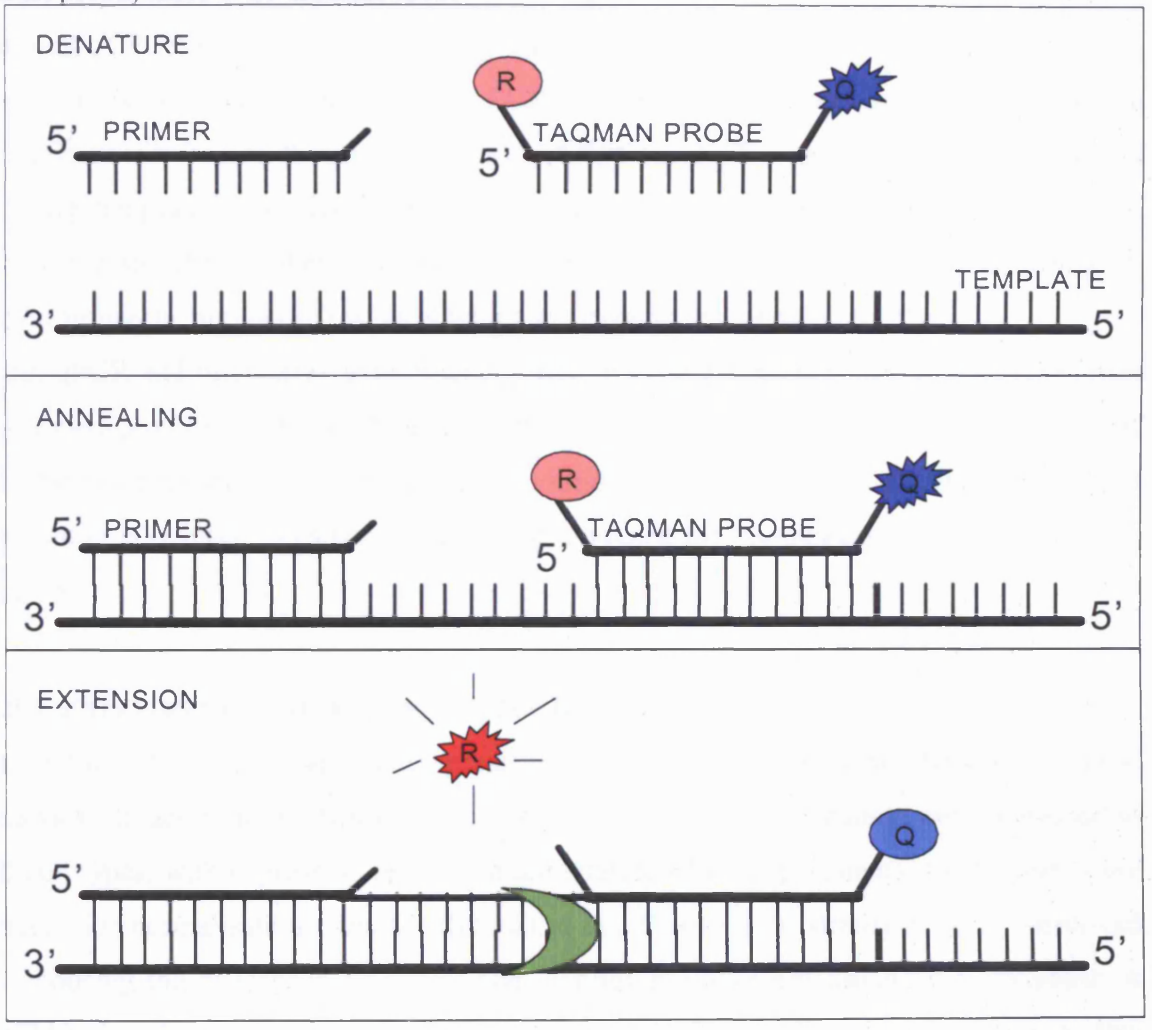


Figure 5.20 Mode of Action of TaqMan® Probe

TaqMan® probes contain a fluorescent reporter dye (R) on the 5' end and a quenching dye (Q) on the 3' base. On irradiation, the excited fluorescent dye transfers energy to the nearby quenching dye by FRET. Due to their close proximity in an intact probe no fluorescence is generated. During PCR annealing the TaqMan® probe binds to its complementary sequence on the DNA template. When the template is copied during the extension phase of a PCR reaction the 5' exonuclease activity of the enzyme cleaves the probe. This frees the reporter dye from the constraints of the quenching dye resulting in reporter fluorescence. In qPCR reactions, as the number of cycles increases the amount of fluorescence increases in proportion to the number of copied molecules. Therefore the accumulation of PCR products is detected by monitoring the increase of fluorescence.

During the annealing stage of PCR amplification the probe binds to it's complementary sequence on the DNA template. On extension, the 5' exonuclease activity of Taq polymerase cleaves the probe, freeing the reporter dye from the FRET activity of the quenching dye (Holland et al., 1991). This results in fluorescence which is detected in a modified thermocycler unit (figure 5.20). As the reaction progresses there is an increase in fluorescence due to further probe cleavage, which is directly proportional to the amount of PCR product. By recording the amount of fluorescence generated after each cycle, it is possible to monitor the PCR reaction, at the exponential phase, where significant increases in the amount of product first occur. The greater the copy number of starting template molecules in a sample, the faster this exponential phase is reached. Thus, the specific number of transcript molecules in any samples can be compared by determining the number of cycles at which the exponential phase is reached.

Both qPCR and microarray experiments measure gene expression, though the techniques are fundamentally different. Numerous studies have used qPCR to verify the results of microarray experiments, as the technical variation, associated with the use of a single technique is reduced, thus providing strong evidence for the observed data (Rajeevan et al., 2001).

5.5.1 EXPERIMENTAL APPROACH FOR QPCR ANALYSIS

From the list of significant genes presented in table 5.6 three were chosen for qPCR analysis. In addition two housekeeping genes, proposed to be constitutively expressed in all cell types, with expression levels that are unaffected by experimental treatments, were chosen for normalisation purposes (Suzuki et al., 2000). Two strategies were employed in choosing the housekeeping genes; one, Peptidylprolyl isomerase A or Cyclophilin A (PPIA), has been employed in many gene expression studies for normalising data (Alfonso et al. 2004), the other, vacuolar protein sorting protein 4a (Vps4a) was chosen on the basis that its expression levels were constant in the microarray experiments conducted here.

This was determined by calculating the co-efficient of variation (CV = standard deviation / mean) for the $\log_2$ (Exp/Ref) ratios for every gene, amongst all samples used in these experiments. Low CV values are indicative of genes stable to treatment effects. Vps 4a had a CV of 0.08, placing it in a group of genes, approximately 5% of the total number, with the lowest CV.

Apart from PPIA, the NIA ID of the significant clones (table 5.7) was used to obtain the GenBank Accession number (available in Appendix 4 and presented in table 5.7). The GenBank number provided clone sequence information which enabled a BLAST search to be carried out to obtain the identity of genes, the sequences of which showed significant homology to the GenBank sequence (http://www.ncbi.nlm.nih.gov/BLAST/). The name of each gene is provided in table 5.4 in addition to a LocusLink ID. Using the LocusLink database (http://www.ncbi.nlm.nih.gov/LocusLink/), full length mRNA sequence was obtained and aligned to the genomic clone for each gene to determine intron / exon junctions within the mRNA sequence. This enabled probes to be designed (section 2.20.1) across the junctions, which, importantly, ensured that any genomic DNA present in the experimental mRNA samples and amplified during the qPCR stage, would not generate confounding fluorescence.

Calibration standards were created (section 2.20.2) and used to optimise the amplification reaction of each gene, in an ABI Prism® 7700 sequence detection system (section 2.20.4). Each dilution of a series, made from the calibration standards, was added in triplicate to an assay plate. cDNA was generated, using total RNA (section 2.8.7) from the experimental samples derived from barrel cortices subjected to 1 day deprivation (section 5.3.1). The cDNA was diluted and for each sample equal amounts were incorporated in triplicate, together with the standards, onto the assay plate. Specific primers / probes for each gene, together with other components of the reaction mix, were added to individual plates and the reaction carried out using the cycling parameters stated (section 2.20.3).

Following the QPCR reaction, amplification plots, such as those shown in figures 5.21 and 5.23, were created. Here, the measured fluorescence emitted by each sample during each cycle of a reaction (R_n) was plotted against the cycle number.

As described, through the course of a QPCR, the TaqMan® probe is cleaved by Taq polymerase resulting in fluorescence. The amount of fluorescence is directly proportional to the number of copied molecules being generated at each cycle, therefore the amplification plots show the exponential shape indicative of a standard PCR reaction. The plots show an initial lag phase, followed by the exponential phase, in which the greatest increase in PCR generated molecules occurs, until a plateau is reached and the reaction enters the stationary phase (figure 5.21).

NIA CLONE ID	LOCUSLINK ID	GENBANK ACCESSION NO.	GENE NAME
H3097C06	21807	BG071304	*Tgfb1i4*
H3027D10	14694	BG065107	*Gnb2-rs1*
H3020G04	12215	BG064525	*Bsg*
n/a	319427	n/a	*PPIA*
H3131C09	116733	BG074137	*Vps 4a*

Table 5.7 Information on genes used to validate microarray data using QPCR. The NIA clone ID in the table was used to identify genes showing homology to the sequences represented by the NIA clone ID. Using the LocusLink ID intron / exon junctions were identified in the mRNA sequences of each gene. Primers and probes used in QPCR were designed using the 5' end of the mRNA sequence, with probes being designed across intron / exon boundaries.

The number of template molecules in the starting solution affects the number of cycles it took, for a particular sample to reach the exponential phase. This can be clearly seen in figure 5.21, in which more concentrated standards enter the exponential phase before less concentrated ones. Thus, the amount of specific transcript in different samples, can be compared by calculating the number of cycles it takes each sample to reach the exponential phase.

A fluorescence threshold was set, at some point in the exponential phase, and the number of cycles it took for each sample to reach this threshold (Cycle Threshold or C_T) was extrapolated from the amplification plots (figure 5.21). Using the calibrating standards for each gene, the C_T was plotted against the starting amount of plasmid DNA used in the qPCR reaction and a standard curve was constructed (figure 5.22). The expression level for each gene, in each sample was calculated by extrapolating, from a samples C_T value, the corresponding measure on the standard curve. It should be noted that the standard curve used to extrapolate the data should show a high correlation with the observed data (figure 5.22) ie the standard curve should be a straight line with a large linear range.

5.5.2 RESULTS OF QPCR ANALYSIS AND COMPARISON WITH MICROARRAY DATA

The amounts of each gene in the experimental samples were calculated as described (section 5.5.1). The data is presented in figure 5.24, with error bars showing the variation amongst the technical replicates of each sample. In addition, for the proposed housekeeping genes *Ppia* and *Vps 4a*, the coefficient of variation (CV) is given. This was calculated across all experimental samples (using: std dev / mean of amounts of transcript).

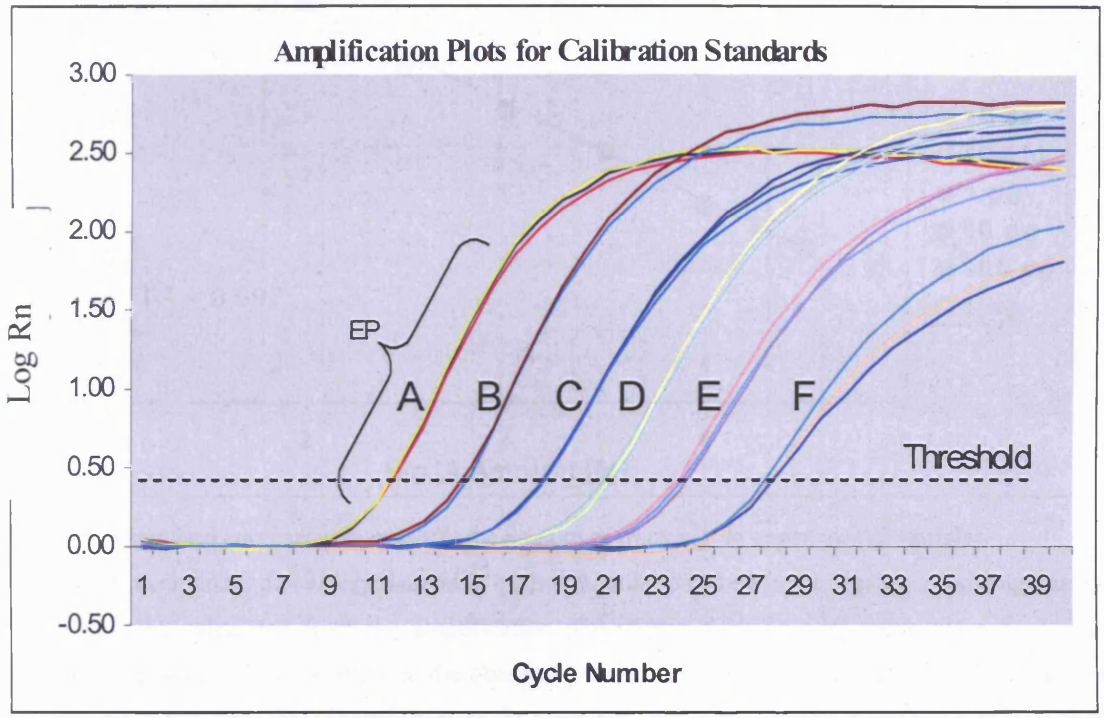


Figure 5.21 The Amplification Plots for each of the calibration standards on a QPCR assay plate

For each calibration standard, the fluorescence measured in the respective well of the assay plate (Rn, y-axis) is plotted as a function of the cycle number. The fluorescence is generated by the degradation of the TaqMan® probe in the reaction mix and is proportional to the amount of PCR product formed. Thus the plots take an exponential shape, typical of the accumulation of product in a standard PCR reaction. The standards were analysed in triplicate and A, B, C, D, E and F are equal to 1 ng, 100 pg, 10 pg, 1 pg, 100 fg and 10 fg starting amounts of plasmid DNA respectively. From the plots it can be seen that standards with higher amounts of starting template start the exponential phase (EP) of amplification at lower cycle numbers. A threshold is set (dotted line) and for every sample, the threshold cycle (CT), the cycle number at which the fluorescence emission exceeds the threshold, is calculated. Thus, standards with higher amounts of starting template have lower CT values.

Vps 4a had the lowest CV and is therefore the most stable gene of the two in the experimental samples assayed. For this reason *Vps 4a* was chosen to normalise the amounts of genes in experimental samples. Briefly, the purpose of normalisation is to remove the systematic errors associated with QPCR. These errors may come from the

relative efficiency of the reverse transcription (RT) reaction used to generate cDNA or the amount of cDNA being added for each sample.

For each gene, the technical replicate measures for each sample, were averaged and normalised by dividing each by the corresponding measurement for *Vps 4a*. The relative

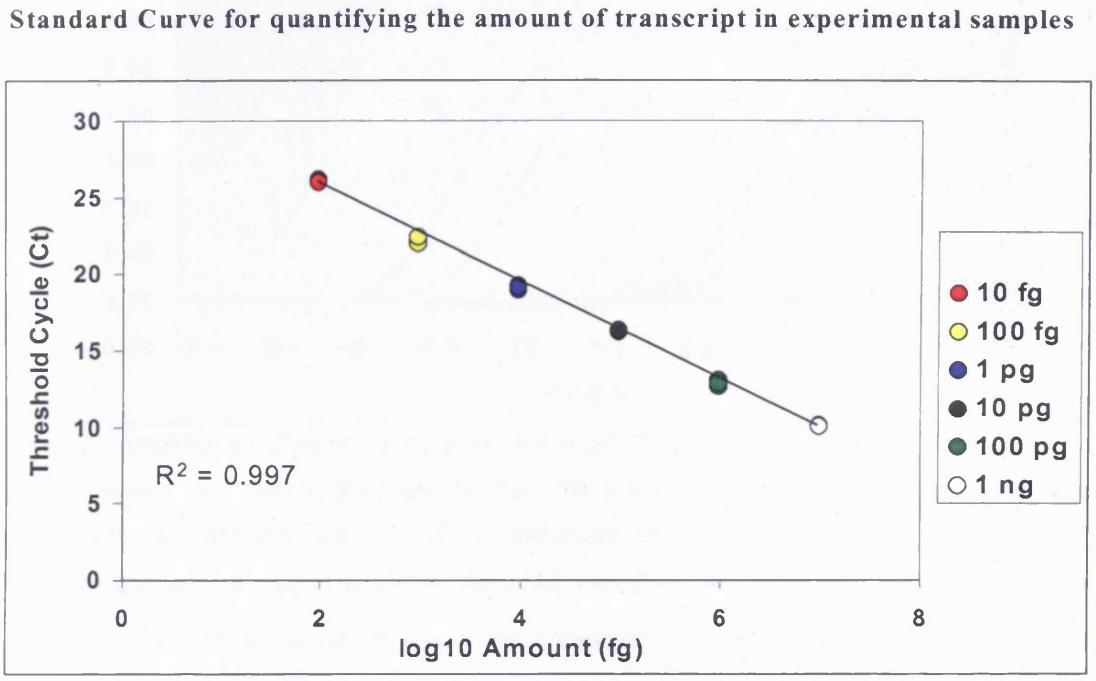


Figure 5.22 A standard curve used to quantify the amount of transcript in experimental samples
The chart plots the amount of starting plasmid in each standard (described in the legend, x-axis) against the CT value (y-axis), calculated from the amplification plot (figure 5.21). An R^2 value close to 1 is an indication of a good fit of the trendline to the observed values. Thus the trendline here reliably describes the relationship in the data. This is important as the trendline, or standard curve, is used as a reference for extrapolating quantitative information about the experimental samples used in the assay.

transcript levels, averaged across the treatment applied to the samples (either ALL, CB or DEP) are provided in figure 5.24 with error bars representing the standard error. The same data from the microarray experiments is also provided in figure 5.25. Given that all genes, apart from *Vps4a*, were declared significantly different amongst the three treatment groups following the microarray experiments, statistical testing was applied to the QPCR data.

Firstly, the relative measures for each sample, were tested gene-by-gene for normality using the Ryan-Joiner method previously described (section 2.20.5). The data for each gene was normally distributed ($p>0.01$), therefore Student's t-tests were carried out to identify any genes significant amongst the treatment groups. Three tests were carried out

for each gene; CB v ALL, CB v DEP, and ALL v DEP and the results are presented in table 5.8.

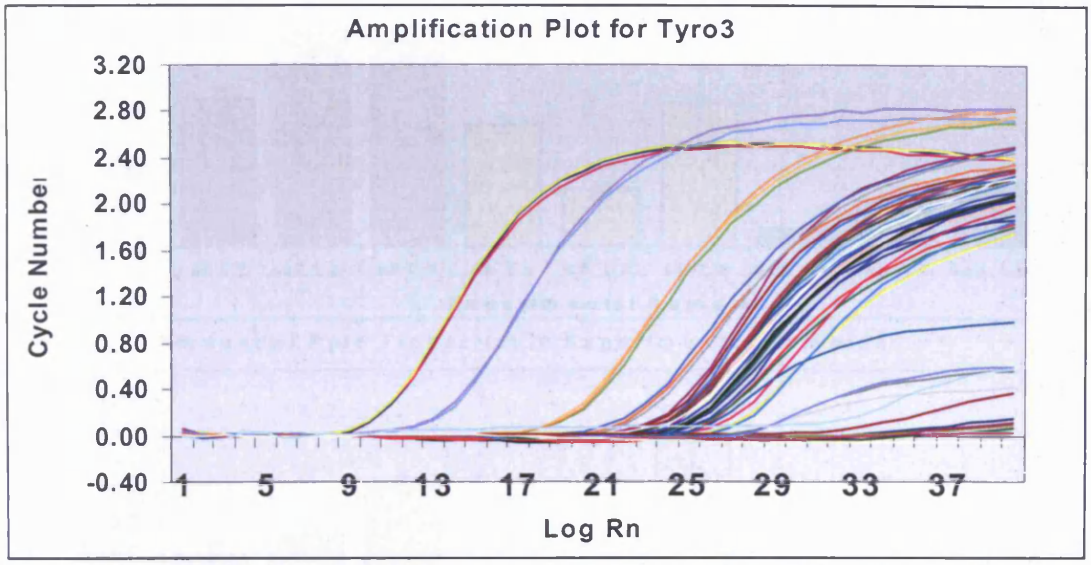


Figure 5.23 Amplification Plots of the *Bsg* gene used in QPCR validation of Microarray Results

The amplification plots used to quantitate the transcript levels of the *Bsg* gene in experimental samples. This plot is the same as that shown in figure 5.19 with experimental samples added.

The three transcripts from the genes *Tgfb1i4*, *Gnb2 rs1* and *Bsg* all show some levels of significant difference, amongst the three treatment groups. The first of these genes *Tgfb1i4* has a QPCR expression profile that indeed resembles that from the microarray data. Here, the transcript appears upregulated in CB treatment compared to both DEP and ALL treatments (figure 5.25). Though the p-value of CB v DEP is not significant (0.085), it is close enough to suggest that a trend does exist. It is possible that the relatively large error bars may hide true significance which would become apparent if more biological replicates were carried out.

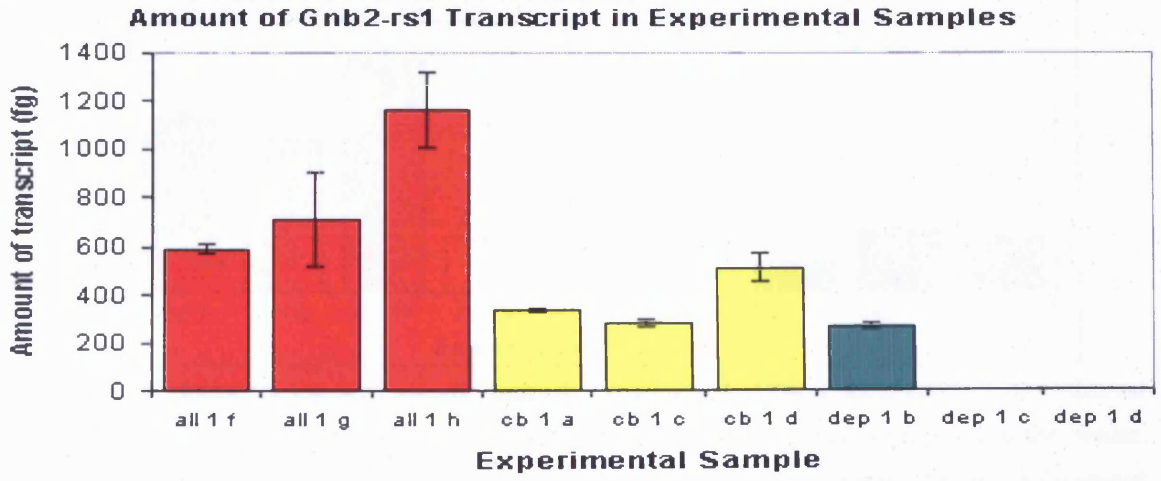


Figure 5.24 (continued overleaf)

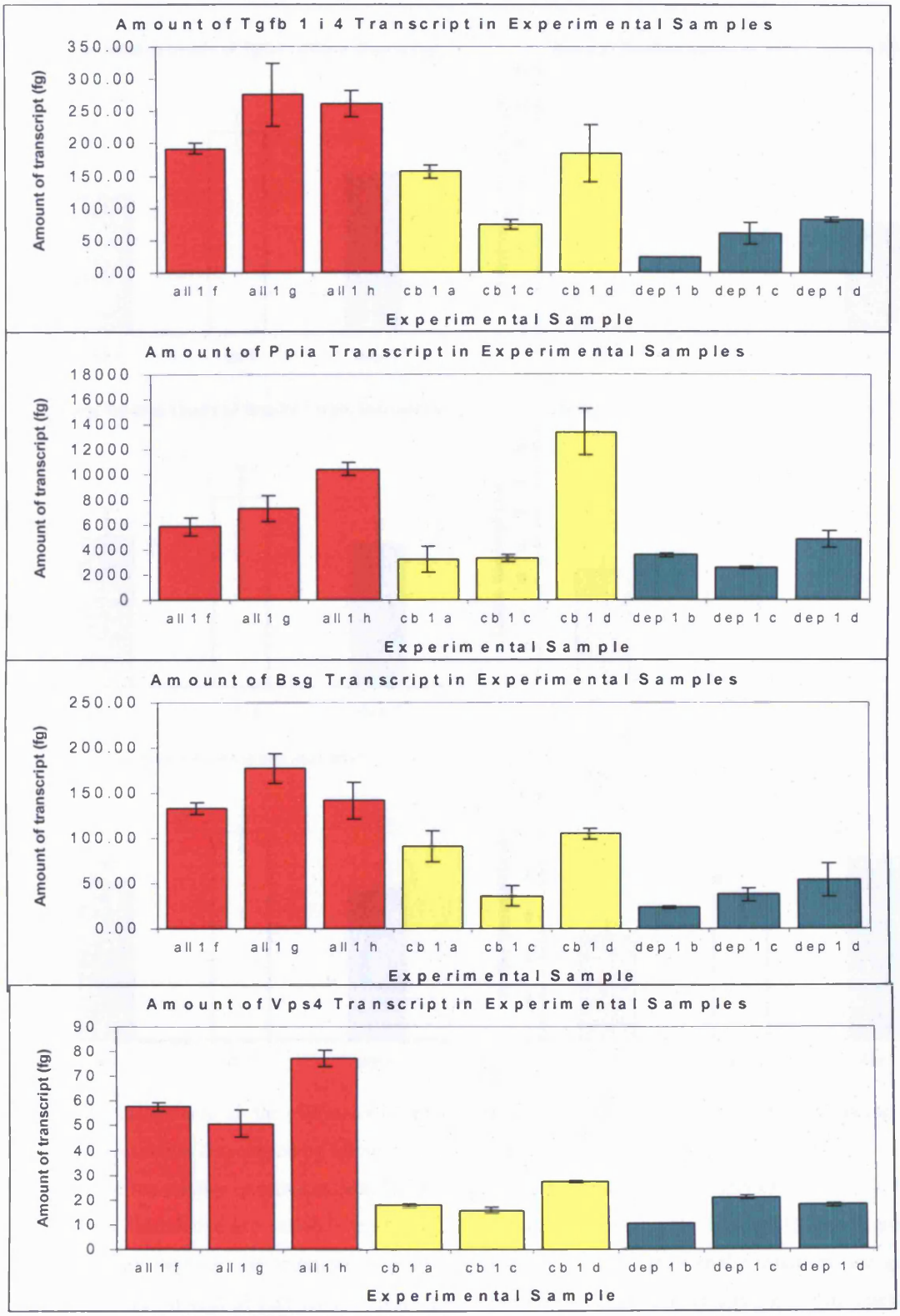


Figure 5.24 The transcript levels of specific genes in experimental samples subjected to 1 day vibrissal deprivation, as determined by QPCR. In each chart the identity of the genes is provided in the header. Each sample is identified by the labels on the x-axis. The y-axis provides the transcript levels (averaged over three replicates) of the gene, in fg, within each sample. The treatments applied to the vibrissae of the animals are colour coded; ALL-RED, CB-YELLOW and DEP-GREEN.

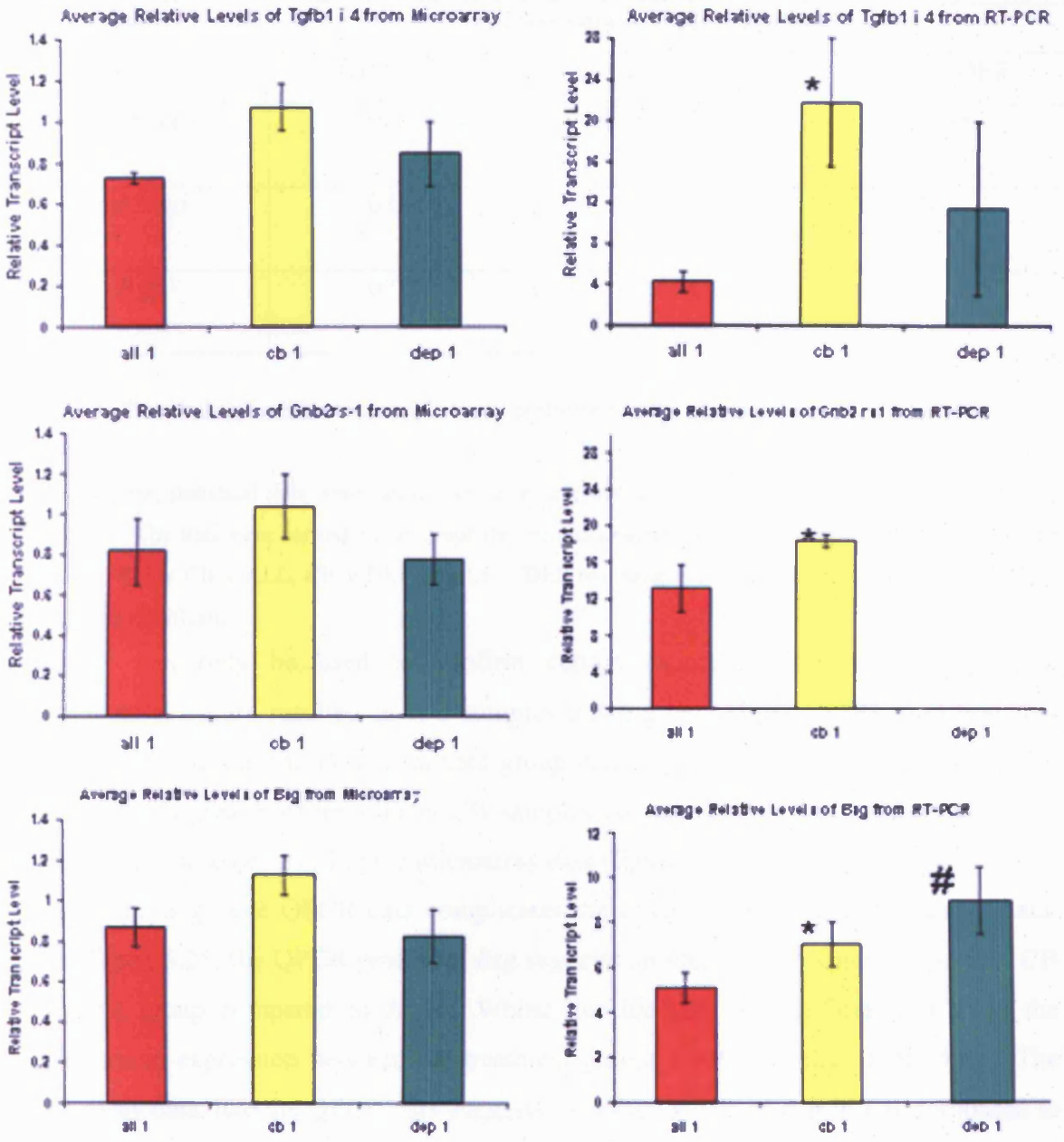


Figure 5.25 Comparison of the expression levels of transcripts in the mouse barrel cortex following 1 day whisker deprivation as determined by Microarray and RT-PCR

For each gene the relative expression data, from both microarray and RT-PCR analysis are provided. For both methods, the relative expression levels (to reference sample for microarray data and to Vps 4a gene for RT-PCR) are averaged across technical replicates for microarray data and over the treatment assigned to three independent biological replicates. Error bars are provided. Note that in this figure the microarray expression ratios are unlogged, allowing standard error to be calculated.

* indicates a p-value < 0.05 for a two-class, one-tailed t-test comparing CB treatment with ALL.

+ indicates a p-value < 0.05 for a two-class, one-tailed t-test comparing CB treatment with DEP.

# indicates a p-value < 0.05 for a two-class, one-tailed t-test comparing ALL treatment with DEP.

Gene	Two sample t-test performed		
	CB v ALL	CB v DEP	ALL v DEP
Tgfb1i4	0.019	0.085	0.138
Gnb2-rs1	0.030	NA	NA
Bsg	0.027	0.074	0.015

Table 5.8 The results of various two sample t-tests performed on the expression levels of genes determined by QPCR.

For each gene, statistical tests were carried out on the expression levels, relative to *Vps 4a*, as determined by QPCR. The tests were carried out amongst the various treatments (CB, ALL and DEP) applied to the samples. Either CB v ALL, CB v DEP or ALL v DEP two-sample t-tests were carried out with $p<0.05$ considered significant.

Gnb2rs1 can only be used to confirm certain aspects of the microarray data. Experimental failure resulted in two samples missing data (figure 5.24), meaning that statistical tests using the DEP treatment group could not be carried out. However, the trend towards greater expression in CB samples compared with ALL, observed in the QPCR data is also observed in the microarray data (figure 5.25).

Finally, the *Bsg* gene QPCR data complicates the confirmation of the microarray data. From figure 5.25, the QPCR profile of *Bsg* suggests an increase in expression in the CB treatment group compared to ALL. Whilst this increase is significant, so to is the difference in expression between the treatment groups DEP and ALL (table 5.8). The microarray data, like the QPCR data suggests increased levels of *Bsg* in CB compared to ALL. However, whereas the microarray data shows a similar increase in CB compared to DEP, the QPCR data shows the opposite trend ie there is an almost significant (p-value 0.074) increase in *Bsg* levels in the DEP treatment group compared to CB.

5.6 DISCUSSION

This chapter used the same basic concepts, employed in chapter 4, to discover genes differentially expressed in the mouse barrel cortex following vibrissal deprivation. The overall aim was to identify genes regulated when the vibrissae are deprived in a manner that induces neuronal plasticity within the system. Using microarray technology transcript levels in the barrel cortex were compared following three vibrissal treatments;

Chessboard (CB), which invokes plasticity in the barrel cortex, no deprivation (ALL) and all vibrissae deprived (DEP). This ensured that genes differentially expressed and associated with plasticity, could be differentiated from genes whose expression levels change in response to changes in activity levels in the system. This is akin to experiments that have studied changes in gene expression associated with mouse visual cortex plasticity (Mataga et al., 2001).

In attempting to compare three treatments, in an experiment which only allows two comparisons, the design is an incomplete block (Cochran and Cox, 1951). In order to choose a suitable experimental design the allocations of replication in the experiments was first studied. In order to make assumptions, regarding differential gene expression, at the population level three biological replicates of each treatment, at each time point, were required. This was facilitated through the use of an in vitro-transcription reaction (Van Gelder et al., 1990) thoroughly optimised in section 3.9. In using this reaction, the pooling of experimental samples required in Chapter 4 was not required and a substantial increase in the power, with which differential expression could be statistically tested, was afforded. Also, to reduce the impact of technical error, replicated measurements were to be taken for each experimental sample (section 5.3.1). Experimenters are still restricted by the relatively high cost of carrying out microarray hybridisations, therefore, resource allocations required for the loop and reference design, using the number of replicates and timepoints stated were studied. Users of microarray generally use two experimental designs to overcome this problem; the loop and reference. In order to carry out this series of experiments the reference design was shown to be the more cost-effective of the two (section 5.2.2). In opting for the reference design in this series of experiments, not only have resources been saved but also the impact of experimental failure (approximately 10%, figure 5.13) has been reduced eg time point 4 days analysis was still possible. The reference design offers a simplicity and robustness that is not offered by the loop design. Experimental failure has far more impact in the loop design compared to the reference design, as the loss of as few as one or two hybridisation in a loop, can compromise the analysis of an entire series of experiments (Yang and Speed, 2002, Dobbin and Simon, 2002). In contrast the reference design is relatively stable to experimental failure with the only impact being a loss of statistical power. The reference design has found a niche in class prediction studies, where global gene expression profiling is proposed as a method for the diagnosis of disease (Wang et al., 2003, Golub et al. 1999).

All hybridisations carried out in this chapter are shown in the design plan (figure 5.6). A potential source of false positives in microarray experiments is dye bias that has been observed for certain transcripts (figure 4.13) and is the source of much discussion in the microarray field (Tseng et al., 2001, Kerr et al., 2001). In order to effectively remove its effects biological samples were assayed twice with dye allocations reversed in a technical replication or dye swap.

For each time point in the deprivation series carried out in this chapter (0.4, 1, 2, 4 and 8 days), genes significantly differentially expressed amongst the treatment groups were identified using SAM (Significance Analysis of Microarray, Tusher et al., 2001). SAM was used as the data from hybridisations were not normally distributed (section 5.4.1), meaning traditional parametric statistical tests could not be applied to the data. In SAM the data itself is used through permutations to assign a level of significance, therefore no prior knowledge of the distribution was required. SAM analysis revealed large numbers of transcripts that were differentially expressed however a high percentage of these failed to be replicated when the dye swap data was compared. This revealed the importance of technical replication, but also highlighted the large numbers of false positives generated in these experiments. One theory for the large number of significant genes that fail to be replicated is that the numbers of biological replicates, used in these experiments, provide too few permutations for SAM to accurately estimate the FDR. In the two-class SAM analyses carried out (section 5.4.1) there are 720 possible distinct permutations amongst the labels (9! / 6!). Some research has suggested that this is not enough to accurately construct a good re-sampling distribution (Slonim, 2002, Pan, 2002) and that a minimum of 10 replicates should be used to generate sufficient permutations. That said SAM has been successfully applied to studies with limited sample size (Xiao et al., 2002). Further to this, attempts to validate the microarray data using QPCR suggest that even transcripts that are differentially expressed in both technical replicates cannot be confirmed using this second assay method. For one of these, *Tgfb1i4*, the relative expression levels, as measured by QPCR mirrored the expression profiles, amongst the treatments, of both technical replicates in the microarray experiments (figure 5.25). A second gene, *Gnb2rs1*, showed a similar expression profile for the CB vs DEP treatments, again in QPCR and technically replicated microarray data. The ALL treatments continually failed to generate data in the QPCR, therefore this aspect of the expression profile couldn't be compared (figure 5.24). The final gene assayed, *Bsg*, showed a different profile in the QPCR data (figure 5.25). Though the differences between the ALL and CB treatment

showed an upregulation of *Bsg* following CB treatment, as did the microarray data, the QPCR results suggested that *Bsg* was upregulated in the DEP treatment compared to ALL, the opposite of both technical replicates of the microarray data.

Given the lack of correlation between the technical replicates of the microarray experiments, and the QPCR and microarray data, it appears that some results are false positive. The only way to remove them is to increase the number of biological replicates assayed by microarray.

From the cluster analysis of the expression levels of differentially expressed genes, the overwhelming effect of vibrissal deprivation appears to be a downregulation as greater numbers of vibrissae are removed ie the majority of significant genes are downregulated in both the CB and DEP treatments. This implies that their transcription is regulated by activity levels in the barrel cortex. This effect is perhaps most striking in cluster A in figure 5.16 and 5.19, where the downregulation in CB is approximately half that of DEP. This is not surprising if the genes are associated with activity changes as in CB half the whiskers are removed compared to DEP where all vibrissae are deprived. It is thus quite probable that removing half of all vibrissae resulted in half the activity in the barrel cortex.

Though of interest, these findings do not seem to answer the questions addressed in this thesis ie the identification of genes solely expressed as a result of CB deprivation. As CB invokes plasticity in the mouse barrel field and DEP and ALL do not, the expression profile of genes that play a role in the phenomenon would be expected to show a marked difference to both DEP and ALL. In this series of experiments only one such cluster is observed (figure 5.16, Cluster B). These genes show a distinct upregulation in CB samples compared to both ALL and DEP following one day vibrissal deprivation. This cluster of genes will be described more fully in Chapter 6 where they will be functionally annotated.

CHAPTER SIX

FUNCTIONAL ANNOTATION OF GENES DIFFERENTIALLY EXPRESSED IN RESPONSE TO VIBRISSAL DEPRIVATION

6.1 AIM

The aim of this chapter is to gain an understanding of the function of genes, differentially expressed in the mouse barrel cortex following 1 day of vibrissal deprivation. The genes have been identified by microarray as either plasticity or activity related. It is proposed that a functional annotation will reveal a mechanistic role for the plasticity related genes in experience dependent plasticity in the mouse barrel cortex.

6.2 INTRODUCTION

Microarrays are capable of generating large amounts of data. Conceivably, the entire transcriptome can be compared and contrasted in a single experiment. The output of microarray experiments are typically lists of genes that comprise the transcriptional element of the biological response to experimental and natural perturbations (Schena et al., 1995, DeRisi et al., 1996). The microarray process is still relatively new; for example, concepts of experimental design and statistical analysis of the data, have only recently been recognised (Churchill, 2002). Through statistical testing, observed results can be further substantiated, providing more confidence in the lists of genes generated.

However the overall aim of a microarray experiment goes beyond generating a list of statistically significant genes. Experimenters, generally, wish to apply some sort of functional significance to the observed data in order to understand how these genes contribute to the biological effects under investigation. This information can only be ascertained if the function of the genes under scrutiny is understood.

Previously, investigators have had to rely on the one-at-a-time approach to functionally annotate gene lists, typically through literature searches. While this approach is entirely adequate for small numbers of genes, the large amounts of data generated in microarray experiments means this often becomes the rate-limiting step. To this end software packages have been developed which attempt to increase the rate at which microarray data can be functionally analysed.

In this chapter two such software packages will be used in an attempt to assign functional significance to the lists of genes, generated in chapter 5, that appear significantly differentially expressed in the mouse barrel cortex, following one day vibrissal deprivation (Appendix 10 and table 5.6). Both DAVID (Database for Annotation, Visualisation and Integrated Discovery, Dennis Jr. et al., 2003) and EASE (Expression Analysis Systematic Explorer, Hosack et al., 2003) are accessible via the internet;
http://david.niaid.nih.gov/david/ease.htm
DAVID annotates genes in a particular list according to shared categorical data for Gene Ontology, protein domains and biochemical pathway membership. The software collates the information about a particular gene from many different sources, and uses keywords, which relate to a particular gene's function, to group every gene in a list according to common function. EASE provides theme discovery within a list of genes, by taking the functional annotations obtained through DAVID and identifying any functional group statistically significantly overrepresented, compared to the populations from which they were drawn ie the genes printed on the microarray slide.

6.3 DAVID

6.3.1 OVERVIEW OF DAVID

DAVID is an online tool which uses different sources to collate functional information, literature and sequence information for a list of genes (Dennis Jr. et al., 2003, Wheeler et al., 2000, Rebhan et al., 1998, Kanehisa and Goto, 2000). This information is obtained from databases accessed through the internet and include;

LocusLink - http://www.ncbi.nlm.nih.gov/projects/LocusLink/

GeneCards - http://bioinfo.weizmann.ac.il/cards/index.shtml

Kyoto Encyclopaedia of Genes and Genomes (KEGG) - http://www.genome.jp/kegg/

PubMed Central - (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed).
DAVID provides four main tools for annotating gene lists. Annotation Tool, GO Charts, KEGG Charts and Domain Charts.

Annotation tool

For a list of genes DAVID allows a number of annotation fields listed to be selected. The output is a table in which genes are hyperlinked to external databases from where information relating to a particular gene can be obtained. The fields available are;

GenBank* – Accession number corresponding to the nucleotide sequence
Unigene* – Cluster containing sequences that represent a particular gene
LocusLink* – Unique and stable identifier for curated genetic loci
RefSeq* – Reference sequence standards for mRNA
Gene Symbol – Official gene symbol included in the Locus Report provided by the NCBI
Gene Name - Official gene name included in the Locus Report provided by the NCBI
OMIM* – catalogue of human genes and genetic disorders
Summary – Functional Summaries included in the Locus Report provided by the NCBI
Gene Ontology – Controlled vocabulary applied to the functions of genes and proteins, as in the Locus Report provided by the NCBI
* - Annotation Fields hyperlinked

While all these fields can be accessed through the NCBI website one at a time, DAVID facilitates the annotation of many genes in a single input. This saves on the laborious chore of data mining which is noticeable with large numbers of genes, typical of the output of a microarray experiment.

GO Charts

The Gene Ontology (GO) Consortium (http://www.geneontology.org/) have created a controlled vocabulary and applied a structured language to the functions of genes in all organisms. Using this vocabulary functional information for each gene in an input list can be collated. The GO terms are structured into three categories; Biological Process, Molecular Function and Cellular Component.

The broadest category of the three is Biological Process which is composed of terms such as metabolism, cell differentiation and reproduction, processes in which large groups of molecules, with a common goal, act in concert. Molecular Function describes the actions of individual gene products eg structural molecule activity, antioxidant activity,

transcription regulator activity. The cellular component category is made up of descriptive terms relating to subcellular structures, locations and macromolecular complexes including microtubule cytoskeleton, large ribosomal unit and cell-matrix junction.

After choosing a category the user is able to specify a level of coverage which either selects more genes and fewer terms (low specificity), or fewer genes and more terms (high specificity).

The output of the analysis for each category is a bar chart (figure 6.1), in which the categories are ordered; those containing the largest number of gene terms are at the top and those with the fewest at the bottom. Thus the distributions of functions amongst genes in the input list can be visualised.

KEGG Charts

The KEGG (Kyoto Encyclopaedia of Genes and Genomes) Charts use the same underlying principles of GO Charts ie a functional term is assigned to each gene. Where the terms for GO Charts relate to the function of a gene the KEGG term describes a gene's involvement in a particular biochemical pathway. The KEGG terms are derived from the KEGG consortium and are also applicable across all organisms (http://www.genome.jp/KEGG/), examples include ATP synthesis and purine metabolism. Each pathway is hyperlinked to the KEGG pathway diagram where the gene from the input list is highlighted. Again the overall aim is to identify those particular pathways that include genes from the input list.

Domain Chart

The aim in these analyses is to display the distribution of genes, from input lists, among PFAM protein domains. This function essentially groups genes on the basis of similarites in domains, contained in the 3D structure, of their functional proteins, examples including RhoGAP and WD40. The domain designations are linked to the Conserved Domain Database at the NCBI
(http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=cdd).

6.3.2 FUNCTIONAL ANNOTATION OF THE NIA 15K CLONE SET

Functional interpretation requires that published information regarding genes in a list is available. Different databases, especially NCBI Unigene and LocusLink, use various IDs (provided in Appendix 4) to describe the same NIA clones used to print the microarrays used in Chapter 5. These two IDs can be used interchangeably in the NCBI database though it was uncertain which to use in DAVID. In particular it was necessary to identify which ID would provide the most functional annotation. To this end a functional annotation of the entire NIA 15K clone set was conducted using both IDs.

The results from DAVID analysis of the NIA 15K clone set are provided in table 6.1. Approximately 13,000 clones have an associated NCBI Unigene compared with 8,740 clones that have an associated LocusLink. Prior to being imported into DAVID both lists were filtered for replicates. This effectively reduced the number of clones by 17% for the NCBI Unigene IDs and 23% for the LocusLink IDs. This is an important point as it suggests that a certain degree of redundancy is still present in the clone set, despite claims by the creators that almost all clones are unique (Tanaka et al., 2000). While this redundancy may be useful in some situations, such as data verification (section 6.3.3) it does limit the range of the microarray.

Of the filtered genes the LocusLink IDs provide the most functionally annotated data (table 6.1). For each GO Chart category the numbers of clones that have annotated data when the LocusLink IDs are used is more than twice that of the NCBI Unigene IDs.

From table 6.1, 10,745 clones on the NIA clone set have a Unigene ID compared to 6,779 which have a LocusLink ID. Unigene is a system that organises sequences, deposited in the public sequence repository GenBank, into non-redundant gene-oriented clusters (Wheeler et al., 2000). LocusLink attempts to collate the data from NCBI Unigene with other databases to provide descriptive information for genes (Pruitt and Maglott, 2001). The NIA clone set was derived from expressed sequence tags which were deposited in GenBank (Ko et al., 2000), therefore the majority of clones would have a NCBI Unigene ID. The time required for manual curation of the LocusLink database is probably the reason why fewer NIA clones have an associated LocusLink ID compared with Unigene IDs.

One further point to make is that from 15,247 clones in the NIA 15K clone set only approximately 20% provide any functional information. This seems quite a low percentage and will be discussed further in chapter 7.

Number of Genes	NCBI Unigene ID	LocusLink ID
Unfiltered	12808	8740
Filtered for replicates	10745	6779
Identified Biological Process	1827 (17%)	2968 (44%)
Identified Molecular Function	1641 (15%)	2655 (39%)
Identified Cellular Component	1645 (15%)	2678 (30%)

Table 6.1 A Functional annotation of the NIA 15K clone set.
Each clone in the NIA clone set has either, or both, an NCBI Unigene and LocusLink ID. The numbers of clones with either ID are provided in row 2 and after filtering for replicates in row 3. In order to identify which ID would provide the most functional data both lists were subjected to a DAVID analysis. The amount of clones, for which functional information was available within DAVID for three categories of annotation, are provided in rows 4-6, expressed as both the numbers of clones and as a percentage of the number of clones in the initial list.

6.3.3 FUNCTIONAL ANNOTATION OF DIFFERENTIALLY EXPRESSED GENES

Chapter 5 identified two subsets of genes that were significantly differentially expressed in the mouse barrel cortex following one days vibrissal deprivation;

transcripts upregulated in the barrel cortex following 1 days Chessboard deprivation compared to both undeprived and all deprived (figure 5.17, table 5.6) will be referred to here as ***Chessboard upregulated.***

transcripts downregulated in the barrel cortex following Chessboard and total deprivation compared to undeprived (figure 5.16) will be referred to here as ***Any Deprivation downregulated.***

The lists of genes in each cluster are provided in Appendix 10, under the file names Cy3Cy5 Ref time point 1 Chessboard Upregulated and Cy3Cy5 Ref time point 1 Any Deprivation downregulated.

Using the LocusLink ids each subset of genes will be functionally analysed using DAVID Every gene in the Chessboard Upregulated subset had a LocusLink ID compared to only 89 of the 243 genes in the Any Deprivation Downregulated cluster (Table 6.2). Similar numbers of genes within each subset were able to be functionally annotated from their LocusLink ID (Table 6.2, rows 5-7) providing approximately 40 genes in each group to work with.

Number of Genes	Cy3Cy5 Ref time point 1 Chessboard Upregulated Genes	Cy3Cy5 Ref time point 1 Any Deprivation Downregulated Genes
Total Number in Subset	102	243
Number with LocusLink	102	89
Filtered for replicates	90	89
Identified Biological Process	45 (54%)	33 (37%)
Identified MolecularFunction	43 (51%)	39 (44%)
Identified Cellular Component	40 (48%)	34 (38%)

Table 6.2 A functional annotation of genes significantly differentially expressed following vibrissal deprivation. The total number of genes in two clusters identified as significantly differentially expressed following 1 day vibrissal deprivation (figure 5.16) are provided in row 2. Row 1 describes the subset to which the genes belong. The numbers of genes with an associated LocusLink ID is provided (row 3). Prior to running DAVID functional annotation, genes replicated in the subsets were identified and removed such that only one replicate was present in the analysed list. The number of genes, for which functional information was available within DAVID, for three categories of annotation are provided in rows 5-7, expressed as both the numbers of clones and as a percentage of the number of clones in the initial list.

Of interest is the presence of five replicate genes in the chessboard upregulated cluster. The identity of these genes is provided in table 6.3 in addition to various database IDs described previously (section 5.3.4). The presence of replicates of the same genes in a single cluster appears to validate the results of the microarrays. Furthermore the expression profile observed for *Tgfb1i4* has been corroborated by QPCR (section 5.5.2).

This section provides a thorough breakdown of the functional distribution of the genes in the Chessboard Upregulated and Any Deprivation Downregulated subsets. The results are composed of four analyses carried out in DAVID. Each analysis functionally groups the genes according to different categories and each will be presented in turn.

NIA 15K Clone ID	Genbank Accession of target sequence	NCBI Unigene ID (Link)	Gene name (Link to GeneCard)	Link to LocusLink (NCBI)	MGI ID (Link to The Jackson Lab)	TIGR Mouse Gene Index ID (Link)
H3023H11	NM_010240.1	Mm.7500	Ftl1	14325	MGI:95589	TC563864
H3126H12	NM_009080.1	Mm.3229	Rpl26	19941	MGI:106022	TC615532
H3097C06	NM_009366.1	Mm.20927	Tgfb1i4	21807	MGI:109127	TC627419
H3126B08	BC010726.1	Mm.9277	Pla2g7	27226	MGI:1351327	TC604849
H3009F03	NM_016750.1	Mm.916	H2afz	51788	MGI:2143989	TC623540

Table 6.3 Replicate Gene IDs

The IDs of five genes which appear as replicates in the same cluster of genes that show upregulation in response to 1 days chessboard vibrissal deprivation. Section 5.3.4 gives a thorough description of each ID.

The results of a Biological Component functional annotation are presented in figure 6.1 in the form of a histogram. Here the genes are grouped according to Biological Component function defined by the Gene Ontology Consortium as a category in which large groups of molecules, with a common goal, act in concert. Terms in this category, as can be seen from figure 6.1 include protein metabolism, cell growth, biosynthesis etc. While the various functional groups are distributed slightly differently amongst the Chessboard Upregulated and Any Deprivation Downregulated subsets, the majority of genes in each subset have identical functions.

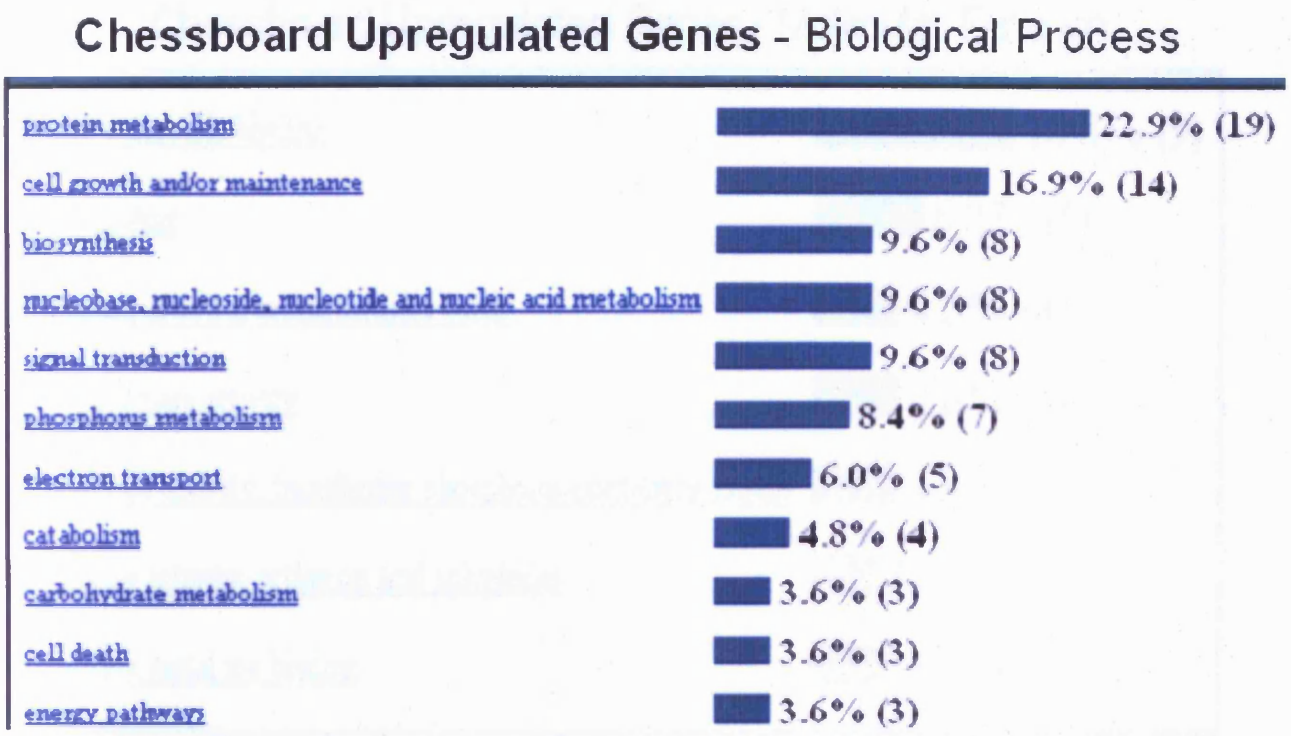
The most heavily populated functional groups, as indicated by the number and percentage of genes provided alongside the groups are;

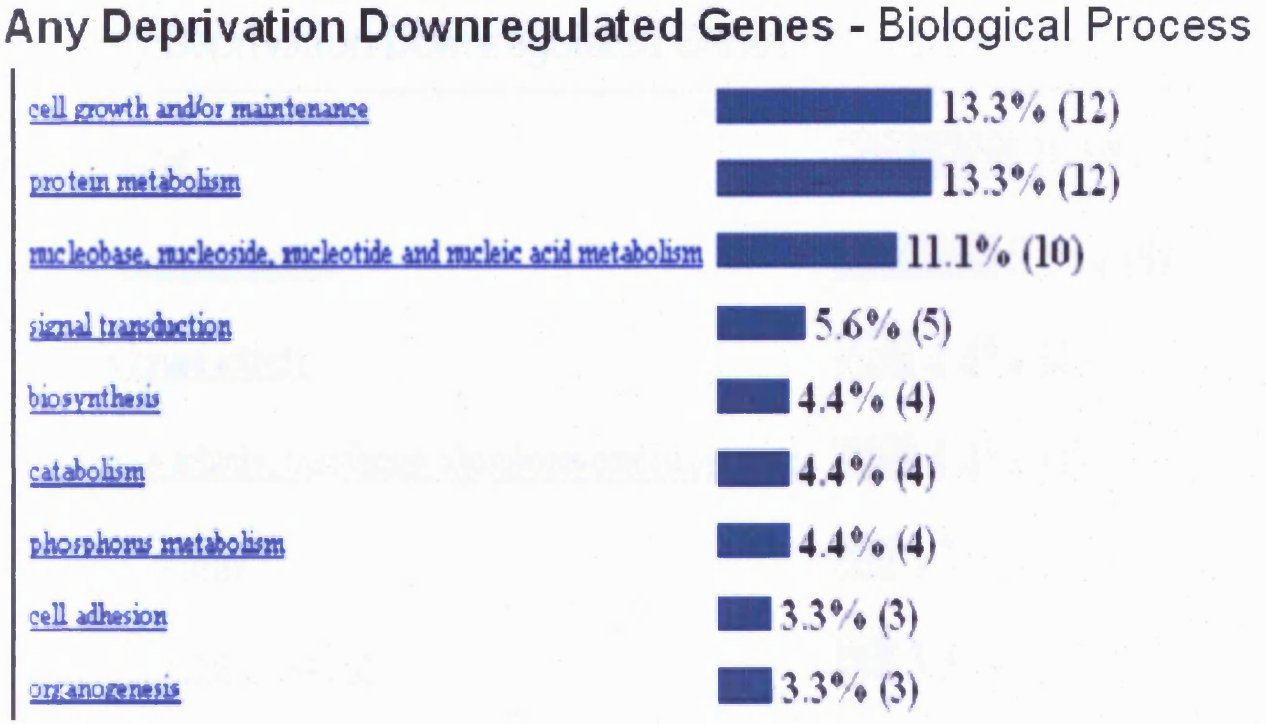
Protein Metabolism, defined as the chemical reactions and physical changes involving a specific protein, rather than proteins in general.

Cell growth and / or maintenance, the processes pertinent to the integrated function of a cell, including cell growth, death and homeostasis.

Nucleobase, nucleoside, nucleotide and nucleic acid metabolism which is an all encompassing term that describes the metabolism of RNA and DNA.

The results of this analysis are shown in figure 6.2. This functional grouping uses terms to describe the actions of individual gene products. For both the Chessboard Upregulated and Any Deprivation Downregulated subsets, the partitioning of function amongst the genes is biased towards purine nucleotide binding and DNA binding, suggesting gene products that bind selectively with purine nucleotides or molecules containing nucleosides or DNA. The functions of the gene products within each subset are almost identical.





The DAVID plot shows the distribution of genes from Chessboard Upregulated and Any Deprivation Downregulated subsets amongst functional categories, defined by the Gene Ontology (GO) Consortium. Each term describes a particular function as within a broad category in which large groups of molecules, with a common goal, act in concert. The identity of each functional category is provided in the charts together with the number of genes that fall into each respective category.

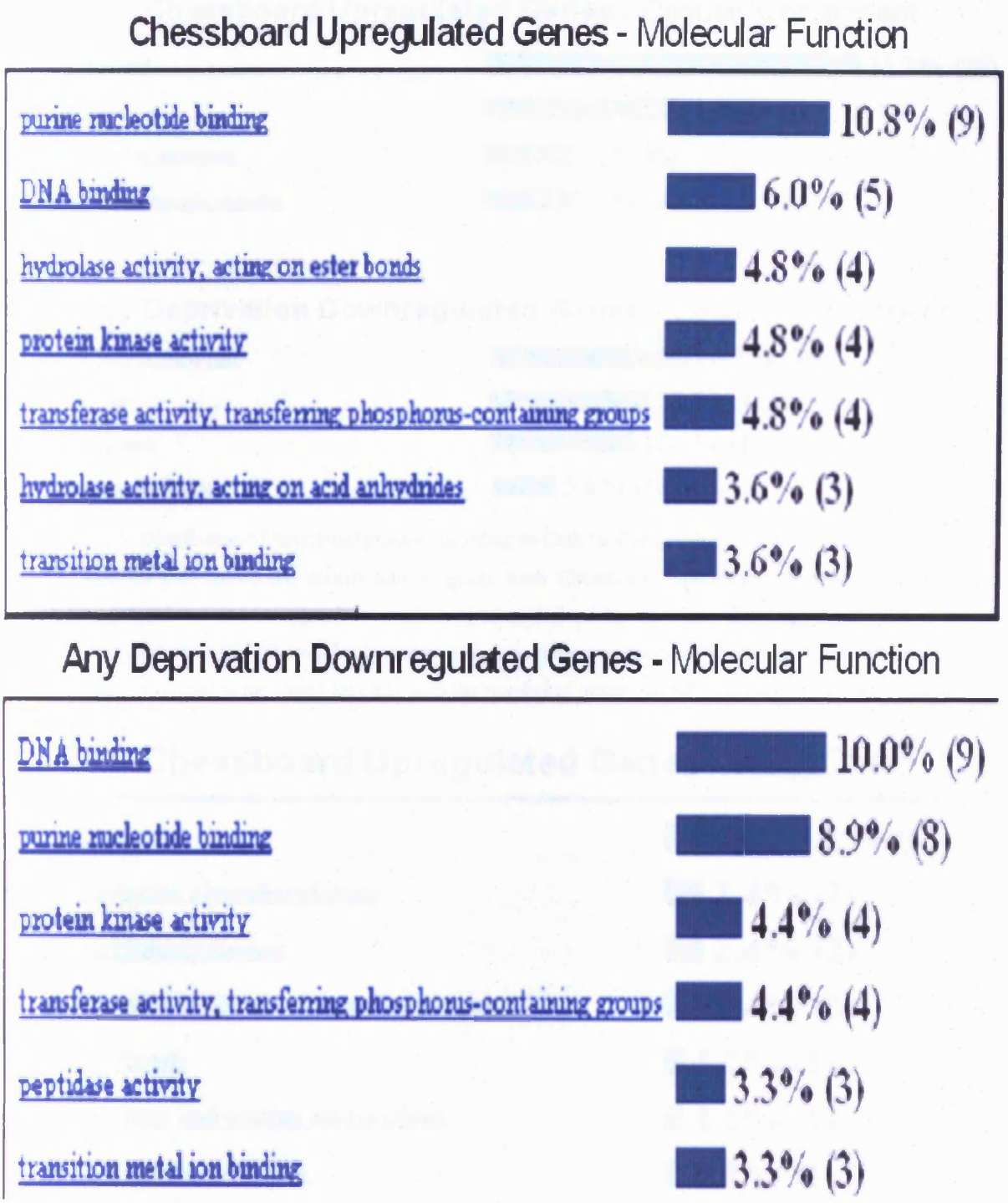


Figure 6.2 Distribution of functional groups according to Molecular Function

The DAVID plot shows the distribution of genes from Chessboard Upregulated and Any Deprivation Downregulated subsets amongst functional categories, defined by the Gene Ontology (GO) Consortium. Each term describes the action of individual gene products. The identity of each functional category is provided in the charts together with the number of genes that fall into each respective category.

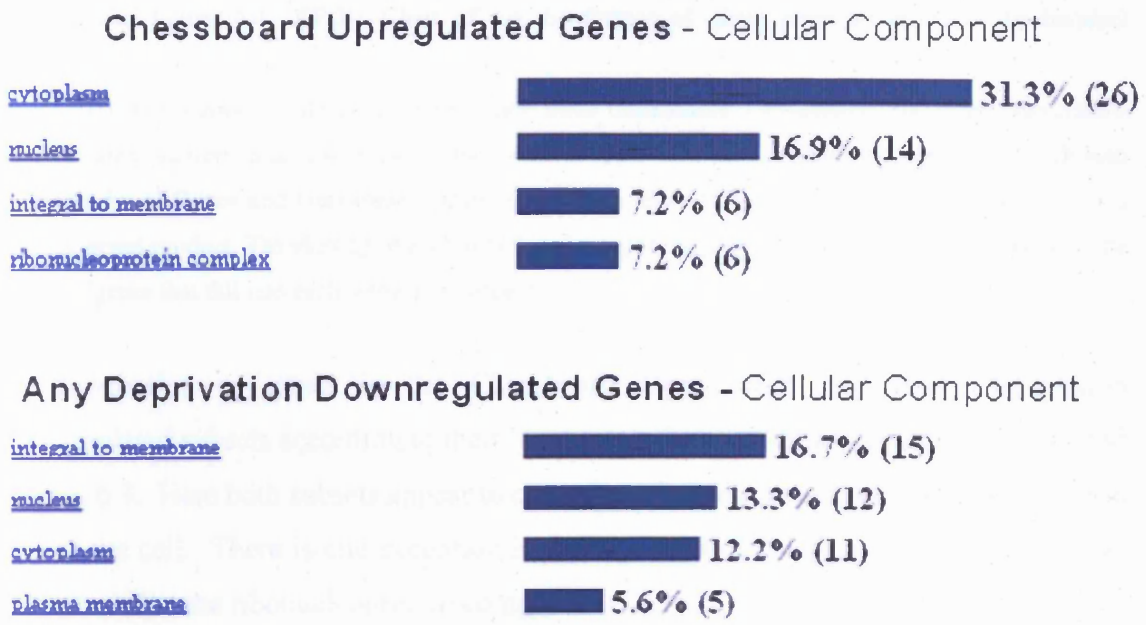
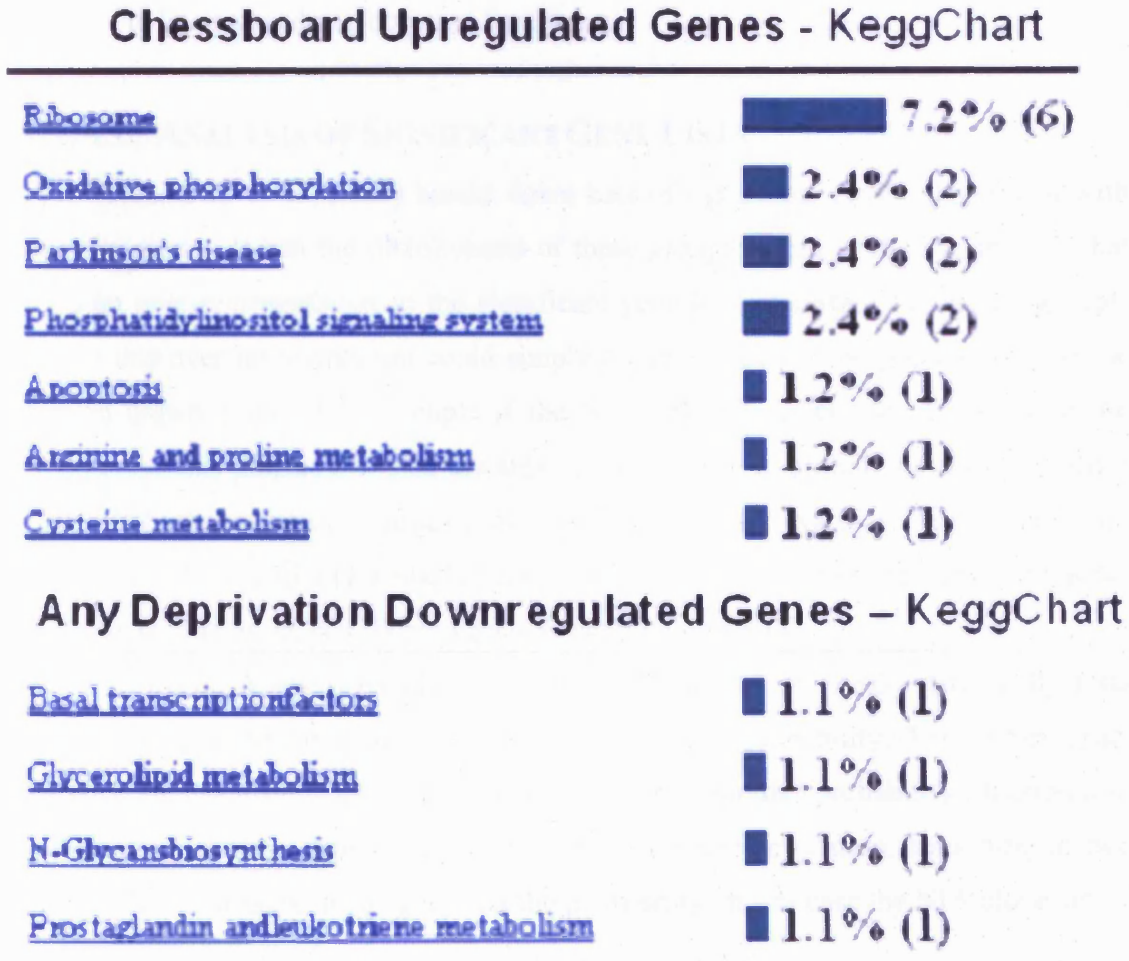


Figure 6.3 Distribution of functional groups according to Cellular Component

The DAVID plot shows the distribution of genes from Chessboard Upregulated and Any Deprivation Downregulated subsets amongst functional categories, defined by the Gene Ontology (GO) Consortium. Each term describes the sub-cellular location of a particular genes product. The identity of each functional category is provided in the charts together with the number of genes that fall into each respective category.



From Overleaf Figure 6.4 KEGG Chart of the distribution of genes amongst amongst Biochemical Pathways

The DAVID plot shows the distribution of genes from Chessboard Upregulated and Any Deprivation Downregulated subsets amongst various biochemical pathways as defined by the KEGG (Kyoto Encyclopaedia of Genes and Genomes) Consortium. Each term describes the sub-cellular location of a particular genes product. The identity of each functional category is provided in the charts together with the number of genes that fall into each respective category.

The distribution of genes in the Chessboard Upregulated and Any Deprivation Downregulated subsets according to their localisation in subcellular structures is provided in figure 6.3. Here both subsets appear to comprise of genes that are localised in the same areas of the cell. There is one exception in that 6 Chessboard Upregulated genes appear to be situated at the ribonucleoprotein complex.

The partitioning of genes in the Chessboard Upregulated and Any Deprivation Downregulated subsets amongst various biochemical pathways, as defined by the Kyoto Encyclopaedia of Genes and Genomes (KEGG) Consortium, is shown in figure 6.4. The numbers in brackets display how many genes in each subset could be assigned to a particular pathway. As can be seen six genes in the Chessboard Upregulated subset show a tendency to be involved in ribosomal functioning.

6.4 EASE ANALYSIS OF SIGNIFICANT GENE LISTS

DAVID (Dennis Jr. et al., 2003) breaks down lists of significant genes into groups with similar functions. When the distributions of these groupings are studied it appears that there is an over-representation in the significant gene lists of certain functional groups. However this over-representation could simply be an artefact of the population that the genes are drawn from. For example if the NIA 15K clone set used to generate the microarrays in this project is biased towards genes involved in protein metabolism then it is more likely that a list of significant genes will be biased towards protein metabolism simply by chance ie will a GO chart of the NIA 15K clone set show the same functional distribution profile as the significant gene GO chart (figure 6.1).

EASE (Expression Analysis Systematic Explorer, Hosack et al., 2003) statistically tests this issue using a variant of the one-tailed Fisher exact probability, The Fisher exact probability is calculated using a Gaussian hypergeometric probability distribution, describing the sampling without replacement from a finite population, consisting of two elements. The population is the genes on the microarray, in this case the NIA clone set,

annotated using the classifications described by DAVID (section 6.3.1). For each classification there are two elements on the array, those that do and those that do not belong to a particular classification. Knowing the number of genes in the population that belong to each element allows the exact probability of randomly sampling a given number of genes, that belong to a particular classification, to be calculated.

The Fisher exact probability is calculated by summing this probability with all probabilities for instances in which there is a greater number of genes which belong within the classification. The EASE score is a conservative adjustment to the one-tailed Fisher exact probability that weights significance in favour of classifications containing more genes. This process, called jackknifing involves removing one gene from a classification and re-calculating the Fisher exact probability. This overcomes situations in which a functional group may have, for example, one representative on the microarray and this gene appears in a significant gene list. From the Fisher exact probability the functional group would be designated as significantly over-represented, however from a global perspective one gene hardly implicates a particular function as being over-represented and if the gene is a false positive the entire functional group is wrong (Hosack et al., 2003).

The LocusLink IDs of the Chessboard Upregulated and Any Deprivation Downregulated subsets of genes, described in section 6.3.3, were imported into the EASE software program (available from http://david.niaid.nih.gov/david/ease.htm). Over-representation of functional groups within each sub-set, was tested by comparison with the LocusLink ID of every clone in the NIA 15K clone set. The results are presented in table 6.4.

The EASE score provides a measure of the probability that a particular functional category is over-represented purely because it is over-represented in the sample from which it was drawn. Thus an EASE < 0.05 represents a p- value < 0.05.

EASE analysis identified three categories of function for both Chessboard Upregulated and Any Deprivation Downregulated subsets of genes. From table 6.4 it can be seen that Chessboard Upregulated gene functions were drawn from three different categories; Molecular, Cellular Component and Biological Component whereas all three functional groups for the Any Deprivation Downregulated genes were from the Cellular Component category. The identities of each sub-category are provided including the GO definition of each. In addition the number of genes in each functional sub-category are provided to assess the how broad the categories are.

For the Chessboard Upregulated genes the list of genes that comprise each functional group are shown in table 6.5. For each gene the particular category it is found in is coded. As can be seen the majority of genes are in the Protein Metabolism group possibly because this group contains the most genes in the GO database. There are six genes that appear in more than one category. In particular the genes Rpl12 and Rpl26 are found in three categories.

Gene Cluster	System	Gene Category	Category %	Category Definition	EASE score
Chessboard Upregulated	GO Molecular Function	structural constituent of ribosome	1.1	The action of a molecule that contributes to the structural integrity of the ribosome	0.03
Chessboard Upregulated	GO Cellular Component	cytosol	1.8	That part of the cytoplasm that does not contain membranous or particulate subcellular components	0.03
Chessboard Upregulated	GO Biological Process	protein metabolism	7.4	The chemical reactions and physical changes involving a specific protein, rather than of proteins in general. Includes protein modification	0.04
Any Deprivation Down-regulated	GO Cellular Component	extracellular space	3.8	That part of a multicellular organism outside the cells proper, usually taken to be outside the plasma membranes, and occupied by fluid	0.02
Any Deprivation Down-regulated	GO Cellular Component	membrane	3.6	The membrane surrounding a cell that separates the cell from its external environment. It consists of a phospholipid bilayer and associated proteins	0.03
Any Deprivation Down-regulated	GO Cellular Component	extracellular	5.3	The space external to the outermost structure of a cell. For cells without external protective or external encapsulating structures this refers to space outside of the plasma membrane. This term covers the host cell environment outside an intracellular parasite	0.03

Table 6.4 Ease Analysis of Functional Groups within Significant Gene Lists

Both Chessboard Upregulated and Any Deprivation Downregulated subsets were analysed for statistically significant over-representation of functional groups. The genes were functionally annotated using the GO system (section 6.3.1) shown in column 2. Each gene was assigned to a particular functional category, assigned using the controlled vocabulary of GO (column 3). The GO definition of each category is provided, in addition to the number of genes currently recognised in that category expressed as a percentage of the total number of genes with annotated function. EASE assigns a level of significance to the numbers of genes in each subset, in a particular category through comparisons with the numbers of genes, from the NIA 15K clone set, in the same category. This is achieved through a modified one-tail Fisher exact probability calculation the EASE score.

LocusLink	Gene Name	Official Gene Symbol	Gene Category		
16828	lactate dehydrogenase 1, A chain	Ldh1	C		
56612	prefoldin 5	Pfdn5	C		PM
18571	programmed cell death 6 interacting protein	Pdcd6ip	C		
19716	reduced expression 3	Rex3	C		
11431	acid phosphatase 1, soluble	Acp1	PM		
11544	ADP-ribosylarginine hydrolase	Adprh	PM		
110355	adrenergic receptor kinase, beta 1	Adrbk1	PM		
226414	aspartyl-tRNA synthetase	Dars	PM		
14232	FK506 binding protein 8	Fkbp8	PM		
14694	guanine nucleotide binding protein, beta 2	Gnb2-rs1	PM		
26396	mitogen activated protein kinase kinase 2	Map2k2	PM		
18762	protein kinase C, zeta	Prkcz	PM		
19045	protein phosphatase 1, alpha isoform	Ppp1ca	PM		
81910	ribosome binding protein 1	Rrbp1	PM		
59048	RIKEN cDNA 1500002I11 gene	1500002I11Rik	PM		
11652	thymoma viral proto-oncogene 2	Akt2	PM		
22027	tumor rejection antigen gp96	Tra1	PM		
22273	ubiquinol-cytochrome c reductase core protein 1	Uqcrc1	PM		
22223	ubiquitin carboxy-terminal hydrolase L1	Uchl1	PM		
269261	ribosomal protein L12	Rpl12	RIB	C	PM
19941	ribosomal protein L26	Rpl26	RIB	C	PM
20054	ribosomal protein S15	Rps15	RIB	C	
67186	ribosomal protein, large P2	Rplp2	RIB		PM
67097	RIKEN cDNA 2210402A09 gene	2210402A09Rik	RIB		
76808	RIKEN cDNA 2510019J09 gene	2510019J09Rik	RIB		PM

Table 6.5 Genes in over-represented functional groups

A list of genes from Chessboard Upregulated subset is provided and the functional group they belong to. The genes comprise the over-represented functional groups identified by EASE. RIB represents structural constituent of ribosome, C represents Cytosol, PM represents protein metabolism.

6.5 LITERATURE SEARCH

DAVID software allows for the rapid annotation of gene lists. In particular it provides a hyperlink that takes the user to the PubMed website at the NCBI where all available literature for a particular gene is collated in abstract form (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed).

The development of software such as DAVID and EASE to identify common functional pathways in lists of genes is arguably simpler than conducting a literature search. However it was speculated as to whether this ability to rapidly sort genes into functional groups has an effect in that some details of gene function, by necessity would be overlooked.

To this end PubMed was used to study the available literature on the Chessboard Upregulated and Any Deprivation Downregulated subsets, in an attempt to identify genes previously implicated as playing a role in neuronal plasticity. Identified in the Chessboard Upregulated subset were four genes that appeared to share a common

function. Of these, three are constituents of the extracellular matrix, Basigin (*Bsg*), Pleiotrophin (*Ptn*) and Dystroglycan (*Dg*) and have all been implicated to the process of neuronal plasticity. The fourth gene mitogen activated protein kinase kinase 2 is a key constituent of MAP Kinase pathway and has been shown to interact with DG (Lauri et al., 1998, Naruhashi et al., 1997, Yamada et al., 2001).

6.6 DISCUSSION

The overall aim of this chapter was to assign functional significance to the Chessboard Upregulated and Any Deprivation Downregulated subsets of genes. DAVID, EASE and the literature search all do this but in distinct ways. The information generated through each will be discussed, with reference to the process of neuronal plasticity, in detail in chapter 7. This particular section will conclude this chapter with a short summary of the differences between each method.

In terms of time taken to analyse the gene list DAVID and EASE were comparable. As expected the literature search took significantly longer. In terms of detail, substantial amounts of information are lost when both DAVID and EASE are used. This is highlighted by identifying three genes through a literature search all of which have been directly related to various forms of plasticity. However of the three methods EASE appears to provide the least information. EASE will only select significantly overrepresented gene lists therefore the similarities amongst both Chessboard Upregulated and Any Deprivation Downregulated subsets were not identified.

DAVID analysis applies no statistical testing to its results therefore no confidence can be generated in any interpretations made on the output. However the striking resemblance between the functions of genes that appear to be associated with plasticity and the genes downregulated when vibrissae are deprived (section 6.3.3) merits further attention. For every functional category tested, with the exception of biochemical pathways the functional product of the genes differentially expressed are the same. Though the distribution of the functional categories differ slightly the overriding impression is that both sets of genes have the same function. In other words chessboard deprivation results in both the up- and downregulation of genes with the same function.

The genes are drawn from relatively large groups with an average of approximately 5,000 members each (data from http://www.geneontology.org/) or approximately 2.8% expressed as a percentage of the total number of annotated genes. Although there is no

way of finding the average percentage there is a suggestion that the terms used to describe each functional group are quite vague. Therefore the similarities in function may simply be a result of the fact that the genes in each subset are poorly annotated.

When the Chessboard Upregulated and Any Deprivation Downregulated subsets of genes are analysed in EASE the similarities observed in DAVID do not occur. EASE calculates the probability that the functional groups that the gene subsets cluster occurs purely chance by comparing then to the functional groupings of every gene on the microarray slide. For the Any Deprivation Downregulated subset of genes it is almost impossible to place any functional relevance to the set of genes apart from the fact that they function extracellularly or in the membrane given that the defining term is so broad.

For the Chessboard Upregulated subset of genes, EASE assigns significance to broadest group observed ie protein metabolism. In the GO database 7.4% of all genes fall into this class. As in the EASE analysis for the Any Deprivation Downregulated subsets of genes it would be almost impossible to ascertain any correlation with the possible mechanism by which plasticity operates from this data. However, as genes whose functional molecules contribute to the structure of ribosomes are also overrepresented following chessboard deprivation it may be speculated that an upregulation of translational machinery may be involved in experience dependent plasticity in the mouse barrel cortex.

In order to investigate the function of Chessboard Upregulated genes in greater detail, a literature search was carried out with a view to identifying any relevant to the process of neuronal plasticity. Of the Chessboard Upregulated genes observed here, three have been identified in other studies that have implicated them as plasticity candidate genes. These genes basigin (*Bsg*), pleiotrophin (*Ptn*) and dystroglycan (*Dg*) code for proteins that are constituents of the extracellular matrix (ECM), a structure that is coming under increasing scrutiny for its relevance to the process of neuronal plasticity in a number of paradigms (Berardi et al., 2004, Dityatev and Schachner, 2003).

The extracellular space surrounding the cells of the central nervous system is composed of a matrix of molecules that carry out a variety of functions including, providing a supportive framework for neurons and glia, ionic regulation and homeostasis. Collectively these molecules are referred to as the ECM and interact with receptors illiciting specific intracellular signalling events that drive processes such as cell proliferation, motility, differentiation, neurite outgrowth and synapse stabilisation (Koppe et al., 1997). The molecules of the ECM condense during development to form

perineuronal nets which completely surround neurons (Bruckner et al., 1993). The nets ensheath neurons and interdigitate and surround synapses.

Dystroglycan (*Dg*) acts to link the ECM with the cytoskeleton. In the central nervous system, DG has been localised to the hippocampus and cerebellar cortex where it may provide a scaffolding for certain synapses (Zaccaria et al., 2001). A role for DG in learning and plasticity was demonstrated by the finding that deletion of *Dg* results in severe deficits in LTP, despite normal neurotransmission in CA1 of the hippocampus. Furthermore deletion of *Dg* did not affect presynaptic neurotransmitter release suggestive of a postsynaptic role (Yamada et al., 2001).

Pleiotrophin (Ptn) is an 18 kDa developmentally regulated protein, initially isolated as a result of studies that attempted to isolate components of the extracellular matrix. Isolation was carried out using heparin affinity fractionation of neurite promoting extracts from young rat brains and fractions were screened using embryonic neurons (Rauvala and Pihlaskari, 1987). *Ptn* is therefore also referred to as *Hb-gam* (Heparin Binding-Growth Associated Molecule). *Hb-gam* is expressed strongly in perinatal rat brain and enhances neurite outgrowth (Rauvala, 1989). In neurons of the developing CNS, HB-GAM is deposited *in vivo* to fiber tracts at the stage of rapid axonal growth, disappearing when axonal development slows. Indeed, both *Hb-gam* mRNA and protein expression, peaks during early post-natal development of the brain (Merenmies and Rauvala, 1990). Expressed in both glial cells and neurons (Wanaka et al., 1993), HB-GAM has been shown to affect postsynaptic differentiation as its application to muscle cells, at the neuromuscular junction, results in the clustering of nicotinic acetylcholine receptors, a marker for membrane differentiation following innervation. Despite a rapid decline in both mRNA and protein levels following CNS development, HB-GAM is still observed in selected populations of neurons, notably pyramidal cells in the CA1 region of the hippocampus (Takeda et al., 1995). Following high-frequency stimulation of the Schaffer collaterals the induced LTP is accompanied by an increase in HB-GAM levels. LTP is inhibited by the application of recombinant HB-GAM which competes with the endogenous protein (Lauri et al., 1998).

The development of 2 transgenic models, that lack and overexpress HB-GAM, have provided further insights into its role in neuronal plasticity (Pavlov et al., 2002). A two-fold over-expression of *Hb-gam* in transgenics is approximately the same level of upregulation observed during LTP. LTP is inhibited to the same degree as when the recombinant protein is added to hippocampal slices. The impairment of LTP is reversed

in *Hb-gam* knockout animals, suggesting that HB-GAM acts as an endogenous, activity induced supressor of LTP (Pavlov et al., 2002).

Basigin (Bsg) is a highly glycosylated transmembrane protein belonging to the immunoglobulin superfamily, thought to play a role in intercellular recognition (Kanekura et al., 1991). Expression has been detected throughout the cerebral cortex with greatest concentration in the pyramidal neurons of layer V (Miyauchi et al., 1990). Mice deficient in *Bsg* show both short term and latent memory impairment suggestive of a neuronal functioning role (Naruhashi et al., 1997).

The three PRGs described, *Dg, Ptn* (*Hb-gam*) and *Bsg* thus appear to affect plasticity in a number of paradigms. It has been proposed that HB-GAM enhances neurite outgrowth through receptors such as N-syndecan and phosphacan which may mediate signal transduction mechanisms that drive neurite outgrowth (Maeda et al., 1996, Kinnunen et al., 1996). It has also been suggested that HB-GAM acts as a postsynaptic extracellular signal that maintains synaptic stability (Takeda et al., 1995). Administration of HB-GAM to culture spinal cord neurons induces clustering of synaptic vesicles which suggested a role in presynaptic differentiation (Dai and Peng, 1998). As HB-GAM can affect both pre- and postsynaptic development it is possible that HB-GAM interferes with the structural modulations that occur at synaptic terminals during LTP (Peng et al., 1995). The precise role of HB-GAM in the mouse barrel cortex is yet to be elucidated, however the fact that it appears to be upregulated under conditions that induce plasticity mark it as a prime candidate plasticity gene.

Bsg has been shown to stimulate fibroblasts to generate matrix metalloproteinases (MMP) (Kataoka et al., 1993). MMPs are a family of enzymes that act extracellularly in the presence of Zn^{2+} and can act to degrade extracellular matrix molecules (Pruijt et al., 1999). Deafferentation of the entorhinal cortex projections to the hippocampus and the subsequent post-lesion sprouting and neuronal plasticity correlate with increased MMP2 activity (Phillips and Reeves, 2001). In hippocampal Dentate gyrus and the neocortex after kainic acid seizures mRNA, protein and enzymatic activity of MMP2 all increase, suggesting a role for the molecule in the synaptic plasticity that follows such events (Szklarczyk et al., 2002, Zhang et al., 1998). Thus the upregulation of *Bsg* observed here may result in an increase of MMP2 in the barrel cortex, and a subsequent degradation of ECM molecules, resulting in altered synaptic stability.

Functionally Bsg is known to play a role in the formation of tenascin-C (TN-C) matrices a process that requires the release of soluble factors by squamous cell carcinoma.

Blocking Bsg function prevents TN-C matrix formation. This could occur through the activation of MMPs as inhibition of their activity directly surpresses TN-C expression in arterial walls and both MMP and TN-C are temporally co-ordinated in calcification processes (Vyavahare et al., 2000, Jian et al., 2001). Thus increased levels of *Bsg*, as observed here could result in an increase in neurite outgrowth mediated through TN-C which could serve to facilitate plasticity in the barrel cortex.

Increased MMP2 could have a direct bearing on the functioning of DG as the protein is a proteolytic target for MMPs (Yamada et al., 2001). It has been observed that DG binds neurexin, a recognition molecule that extends from the presynaptic membrane to the extracellular matrix and interacts with DG. DG extends from postsynaptic space into the neuron where it interacts with postsynaptic dystrophin and hence the postsynaptic cytoskeleton and various molecules involved in signalling pathways. If MPP activity increases during plasticity this could act on DG with both postsynaptic, via dystrophin and cytoskelton, and presynaptic via neurexin conformational changes (Sugita et al., 2001, Gorecki et al., 1998). The implication of this is that DG could be a coincidence detector, a molecule which signals that both the presynaptic and postsynaptic membrane are depolarised. Thus DG could be the signal to neurons that Hebb's Law has been fulfilled (Hebb, 1949). Hebb's Law proposes that neurons whose connection is to be strengthened had to be simultaneously active and DG could be the conduit that results in synaptic strengthening in the barrel cortex.

A wealth of evidence from other studies, discussed here, suggests that basigin (*Bsg*), pleiotrophin (*Ptn*) and dystroglycan (*Dg*) play a key role in a number of plasticity paradigms. Their upregulation in the barrel cortex in response to conditions that induce neuronal plasticity, puts them at the top of the list of candidate genes identified in this study. Whether neurite outgrowth, modulation of synaptic transmission, strengthening of synapses, or a combination of all three processes mediated by these PRGs facilitate plasticity remains to be seen. Attempts have been made to postulate mechanisms whereby the actions of these ECM related PRGs could contribute to plasticity, however it is obvious that their interactions are as complex as the process that results in their upregulation. It is believed that untangling the complex interactions that occur amongst them, and other ECM constituents will provide important information as to the mechanisms that underlie experience dependent plasticity in the mouse barrel cortex.

CHAPTER SEVEN

GENERAL DISCUSSION

7.1 DISCUSSION

Plasticity can be induced by a number of vibrissal deprivation methods, here, the removal of every other vibrissa in a chessboard pattern, induced plasticity across the entire barrel field (Wallace and Fox, 1999). The barrel cortex total RNA was used in a series of microarray experiments in chapters 4 and 5, to identify genes differentially expressed under conditions that induce plasticity.

Plasticity Related Genes (PRG) were hypothesised to have a profile of altered expression levels in chessboard deprived barrel cortices compared to both all deprived barrel fields and undeprived controls. In chapter 5, a sub-set of genes with such an expression profile were identified after 1 days vibrissal deprivation (figure 5.16, table 5.6). These PRGs were upregulated in the chessboard barrel cortex compared to all deprived and undeprived barrel fields. At all other time points studied no further PRG clusters were observed, suggesting that the upregulation of the PRGs identified here occurs rapidly in the temporal scheme of plasticity.

The temporal profile of PRG transcription observed correlates with CRE-mediated gene expression previously studied in transgenic animals. Here mice carrying a *lacZ* reporter gene, downstream of 6 CRE promoters were subjected to vibrissal deprivation over a period of 7 days. Significant upregulation of the reporter gene was observed in plasticity induced cortices following 16 hours of deprivation followed by a rapid drop to basal levels at 7 days (Barth et al., 2000).

The application of high frequency trains of stimuli to the perforant pathway in the hippocampus leads to a sustained enhancement of synaptic transmission that can last for days. This process, long term potentiation (LTP) results in CRE-mediated gene expression, as observed in reporter gene studies, within 2 hours of stimulation. The reporter gene remains upregulated 6 hours post-stimulus (Impey et al., 1996). Investigations using high-throughput analysis have revealed the temporal aspects of LTP associated 'natural' gene expression. Here differential expression of genes occurs 1 hour post-stimulus and many genes are constantly upregulated up to 24 hours post-stimulus, with many transient spikes of differential regulation throughout the time frame (Hong et al., 2004). Thus 'natural' gene expression in l-LTP appears to closely resemble the CRE-mediated transcriptional response in that paradigm.

In the barrel cortex both the reporter gene and 'natural' genes are upregulated only during a very early phase in the induction of plasticity. In contrast increased 'natural' gene expression occurs throughout and beyond l-LTP. LTP has two phases; a short phase associated with immediate early gene expression which lasts for approximately three hours, and a late phase (l-LTP) which is dependent on gene transcription and *de novo* protein synthesis that can last for days (Frey et al., 1988, 1996). In LTP the initial transcriptional response consists of immediate early gene expression, the second transcriptional response that is thought to determine whether potentiation is sustained is dependent on CREB. Given the relatively long period of deprivation required to invoke maximal plasticity in the barrel cortex, it may be the case that the transcriptional response observed in the experiments conducted here and in the CRE-mediated reporter gene study, represent just one wave of transcription. It is possible that this is a transient increase in transcription that is followed by a second more sustained phase of transcription, as in l-LTP. The difference being that in the barrel cortex the initial wave of transcription appears to be CRE-mediated whereas in l-LTP the second more effective wave is CRE-mediated. In order to confirm these hypotheses further experiments are required in which deprivation is carried out for longer periods of time.

Cluster analysis identified Activity Related Gene (ARG) clusters throughout the period of deprivation. The expression profiles revealed that ARGs were downregulated, at all time points in chessboard and all deprived barrel fields compared to undeprived, control levels (figures 5.16 – 5.19). Plotting the numbers of ARGs at each time point allows the temporal effect of vibrissal deprivation over 8 days to be visualised. It appears there is a large initial increase in ARG downregulation following 1 days deprivation, then a return to baseline as the deprivation is maintained (table 5.5). Blanton et al., (2003) have observed similar results when the nerve agent VX has been used to impair the cholinergic nervous system of the hippocampus over a period of 2 weeks. Microarray profiling revealed a significant initial transcriptional response followed by a return to control levels (Blanton et al., 2003).

At the 1 day and 8 day time points there is a well defined pattern to the expression of all the downregulated ARGs. It appears that expression levels decrease as more vibrissae are deprived. Expression of a typical ARG is lowered in chessboard barrel cortices compared to undeprived controls and further lowered when all vibrissae were deprived (figures 5.30 and 5.33). This is consistent with a hypothesis that depriving the vibrissae results in decreased activity levels in the corresponding barrels and this reduction appears to be

associated with a decrease in transcription. Melzer and Steiner (1997) have studied the expression levels of immediate early genes (IEG) in direct respone to vibrissal stimulation. They observed that within 5-15 minutes of stimulation, upregulation of the IEG *zif268* and *c-fos,* occurred in precisely the barrel corresponding to the stimulated vibrissa. Both *zif268* and *c-fos* encode transcription factors that regulate the activity of other genes (Robertson, 1992). Thus a possible explanation for the fact that the majority of ARGs observed here are downregulated is that depriving the vibrissa reduces neuronal activity, blocking expression of *zif268* and *c-fos* which has a direct impact on transcription levels. As the IEG response is confined to stimulated barrels, this would explain why the reduction of expression levels of ARGs in chessboard barrel fields is approximately half that when all vibrissae are deprived, as only half of the vibrissae are removed in a chessboard deprivation.

Attempts were made to confirm the microarray results using qPCR (quantitative PCR). Time constraints limited the number of transcripts assayed using this second method to three; *Tgfb1i4 Gnb2 rs1* and *Bsg*. qPCR analysis validated the differential expression of *Tgfb1i4* and *Gnb2-rs* however *Bsg* showed a very different expression profile between the two methods. These results suggest that the experimental design lacked sufficient power to control the Type II error rate meaning false positive genes are present in the list of PRGs (table 5.6). Type II errors are a consequence of the strengths of microarrays – the ability to measure the expression levels of thousands of transcripts in a single assay eg a microarray spotted with 15,000 transcripts will generate 150 differentially expressed transcripts, at a p-value of 0.01, by chance. Significance Analysis of Microarray (SAM) was used to identify differentially expressed transcripts and uses permutation of the data to calculate the False Discovery Rate (FDR), an estimate of the number of false positives (Tusher et al., 2001). FDRs were kept below 10% for each analysis and experimental samples were assayed twice to account for technical variation. Low numbers of statistically significant transcripts were common to replicate experiments reflecting the inherent variability in microarray experiments (table 5.5). Taken with the fact that the tissue sampled consists of a heterogeneous population of cell types it may have been overly optimistic to use three biological replicates to overcome biological variation and identify truly differentially expressed genes. There is no consensus of opinion as to how many replicates to use in microarray experiments. Through variability modelling Zien et al., (2003) suggested a minimum of 8 replicates and by direct testing Lee et al (2000) found that at least 3 replicates significantly reduces the chances of false positives. With

hindsight it may have been more informative in these experiments to conduct fewer time points with increased sample sizes (possibly 3 time points and 6 biological replicates).

DAVID was used in chapter 6 to functionally annotate the list of PRGs and ARGs identified as differentially expressed following one day vibrissal deprivation (Dennis Jr. et al., 2003). The most significant result of the analysis (section 6.3.3) was the similarity between the functional distributions assigned to each gene list. Both PRGs and ARGs, for almost every analysis carried out, appeared functionally the same. From the biological objectives to which the genes or gene products contribute, through the biochemical activity of the gene products, to the place in the cell where the product is most active, approximately 40% of PRGs and ARGs were functionally very similar. It appears that the transcriptional response, that possibly results in plasticity, is functionally the same as the transcriptional response to stimulation of vibrissae. In other major plasticity paradigms, monocular deprivation in the visual cortex and LTP in the hippocampus, differential gene expression, be it natural (Hong et al., 2004) or a CRE-mediated reporter gene (Pham et al., 1999, Impey et al., 1996) is sustained through and beyond the onset of plasticity. Thus, the situation here, where PRGs are functionally related to ARGs could be due to the fact that only an initial wave of transcription has been observed. This initial wave of transcription could be reflective of increased activity levels in that sub-set of cells that will eventually undergo plastic changes. The system could be 'gearing up' ready for a second wave of transcription. At some later time point, beyond 8 days of deprivation it is possible that additional PRGs will be transcribed with a more plasticity specific function.

According to EASE analysis (section 6.4) (Hosack et al., 2003) both ARGs and PRGs were composed of significantly overrepresented functional groups, though the terms used to define these groups were extremely general eg significantly overrepresented in the ARGs were genes or gene products involved in 'protein metabolism', a functional group with approximately 13,000 members. On average the functional groups identified by EASE, for both ARGs and PRGs consisted of 4,000 members, possibly too large to draw any real conclusions about a possible functional difference between genes related to activity and those related to plasticity as identified here. One interesting result from EASE was the finding that 'structural constituent of ribosome' functional group was overrepresented in the PRGs. This suggests that an upregulation of translational machinery is required for experience dependent plasticity. The requirement for *de novo* protein synthesis is an underlying factor that differentiates short-term and long-lasting

LTP in the hippocampus (l-LTP) (Stanton and Sarvey, 1984, Krug et al., 1984, Frey et al., 1988). The requirement for protein synthesis in barrel cortex plasticity could be investigated using chessboard deprivation, which has proved to be an ideal model to test the effects of various pharmacological agents on neuronal plasticity. The induction of plasticity over the entire area of the barrel cortex, has meant that the localising of various chemicals to the site of plasticity is relatively simple (Wallace and Fox, 1999). To test the requirement for protein synthesis, a blocker such as cyclohexamide infused over the barrel field for a day with chessboard deprivation would be required. A suitable control would be needed, such as an undeprived animal similarly treated, which would account for the fact that blocking protein synthesis could possibly block other key pathways.

Lachance and Chaudhari (2004) have identified 'neuroplasticity', 'neuronal growth' and 'cytoskeletal' related functional groups amongst the genes differentially expressed in the visual cortex when plasticity is invoked through monocular deprivation. Further to this, studies analysing the transcriptional profile of neurons after prolonged NMDA receptor activation have revealed the differential expression of genes clustered into functional categories as diverse as 'cell / adhesion', 'cytoskeleton re-organisation' and 'neurotransmitter related' (Hong et al., 2004) in addition to 'protein folding', 'protein trafficking' and 'protein synthesis'. Taken with the studies of Costigan et al., (2003) who identified genes, associated with re-innervation of dorsal root ganglion, that could be segregated into 16 functional groups of genes, the data presented here suggests the transcriptional response to plasticity in the mouse barrel cortex is composed of genes with not such a diverse range of functions.

It could be possible that the sole purpose of gene expression in the barrel cortex that accompanies plasticity is directed towards protein synthesis. This could mean that the mRNAs that are translated into functional proteins that facilitate plasticity are already present in the neurons. Approximately 250 mRNAs are thought to be localised to dendritic sites, in addition to a whole spectrum of translation initiation and elongation factors (Moccia et al., 2003, Tiedge and Brosius, 1996). The observation that the addition of an artificial mRNA to neurites results in translation of the protein, suggests that local protein synthesis at the synapse can occur independently of the soma (Crino and Eberwine, 1996, Wenzel et al., 1993). It is possible therefore that an upregulation in genes involved in protein synthesis serves to translate the effectors of plasticity that are already present at the synapse eg calcium/calmodulin dependent protein kinase II

(CaMKII), a key plasticity candidate gene (Silva et al., 1992, Hinds et al., 1998, Glazewski et al., 1996).

If it is assumed that the PRGs identified here have a similar functional make-up as ARGs, then a mechanism must exist whereby the functional products of the PRGs act differently in cells that undergo plastic changes. It is known that not all synapses that are made between neurons are potentiated during synaptic plasticity. The enhancement of synaptic transmission is very specific and one rule that governs LTP, *specificity*, states that potentiation of a single synapse will not result in potentiation of all synapses of the same neuron (M^{c}Naughton et al., 1978, Levy and Steward,1979). It could therefore be that the generic response observed here becomes very specific if directed to the correct synapse. One theory, as to how proteins are directly targeted only to those synapses that will be potentiated, is that enhanced synaptic activity results in the presentation of a 'tag'. This tag acts as a memory trace, localised to a specific synapse, sequestering the molecules that facilitate plasticity (Frey and Morris, 1998). Only those synapses that present this tag will be potentiated affording a level of discrimination that allows the synapses of the central nervous system to be selectively enhanced. A possible tag in the sea-slug *Aplysia* is the cytoplasmic polyadenylation binding protein (CPEB) whose mRNA is localised to dendrites. CPEB protein translation increases during enhanced synaptic activity and the molecule undergoes a permanent change in conformation which results in increased translation at the dendrite. Whether this conformational change means CPEB mediated protein translation is enhanced or the sequestering of protein machinery leads to the observed increase is uncertain (Si et al, 2003).

It is possible that a 'second wave of CRE-mediated transcription occurs later in the plasticity time course' that results in the expression of genes, with a range of functions that may be more in keeping with postulated mechanisms for plasticity. This idea is plausible and deserves to be investigated further.

Another reason for the rather generic transcriptional profile that accompanies plasticity, as observed in this work, is that significant changes in expression of other transcripts have been missed ie Type I Errors have been made. As an example QPCR profiling has revealed that stimulation of a single vibrissa results in significant increases in *BDNF* expression in corresponding barrels (Nanda and Mack, 2000). Here *Bdnf* transcript expression cannot be detected above background levels, therefore the findings of Nanda and Mack cannot be replicated. Type I errors may have been made because transcript expression is below the measurable limits of the microarray or the transcripts are

expressed in a large number of cells though differential expression occurs in just a few therefore cellular heterogeneity of the dissected tissue masks true changes.

The vast cellular heterogeneity of the mammalian brain prevails at different levels; different regions have specialized functions reflected in their cellular compositions and within a region several different neuronal cell types may be present. Barth et al., reported Cre lacZ reporter gene expression was confined to approximately 30% of the neurons in layer IV barrels that undergo plasticity (Barth et al., 2000). In this series of experiments RNA was extracted from the entire barrel field and it is possible that only a fraction of the cells dissected are involved in the transcriptional processes that drive plasticity. It may be prudent in future studies to consider techniques that reduce the heterogeneity of the dissected cells and only sample cells that show transcriptional changes associated with plasticity. Indeed the reporter gene positive cells identified by Barth et al (2000) could serve as markers for Cre mediated gene transcription that accompanies plasticity. The findings presented here, that *Tgfb1i4* upregualtion is associated with plasticity means this gene could also serve as a marker for cells that show transcriptional differences in plasticity induced barrel cortices. *In-Situ* hybridisations could be used to label *Tgfb1i4* positive cells or *lacz* processing could be used to identify those cells that undergo *Cre* mediated gene expression. Dissection of the relevant cells could then be achieved by Laser capture microdissection (LCM) which allows for the selective collection of cells of interest from tissue sections (Emmert-Buck et al., 1996). Here, tissue sections are mounted on standard glass slides, and a transparent, 100-mm-thick, ethylene–vinyl acetate film is then placed over the dry section. A laser provides enough energy to transiently melt this thermoplastic film in a precise location, binding it to the targeted cells. The laser diameter can be adjusted from 7.5 to 30μm so that individual cells or a cluster of cells can be selected. Because the plastic film absorbs most of the thermal energy and the pulse lasts for a fraction of a second, little or no detectable damage of biological macromolecules occurs.

After the appropriate cells have been selected, the film and adherent cells are removed, and the unselected tissue remains in contact with the glass slide. These cells can then be subjected to appropriate extraction conditions for molecular analysis. This method has been used in concert with microarrays to identify differential gene expression between groups of neurons within the CNS (Luo et al., 1999) and to study the effects of traumatic brain injury on hippocampal gene expression in specific cell types (Shimamura et al., 2005). Experimenters typically use an RNA amplification technique, tested and used in

this thesis, to generate sufficient aRNA to perform microarray analysis of dissected cells. An alternative approach to reducing cellular heterogeneity could use Fluorescence Activated Cell Sorting (FACS) (Herzenburg et al., 1976). Here cells undergoing Cre mediated gene expression in the barrel cortex could be dissected by replacing the lacZ reporter gene under the control of Cre promoters with a green fluorescent protein (Barth et al., 2000). FACS is a type of flow cytometry, that sorts a suspension of cells into two or more containers, one cell at a time, based upon specific light scattering and fluorescent characteristics of each cell. A suspension of cells from the barrel cortex would be entrained in the center of a narrow, rapidly flowing stream of liquid. GFP positive cells from the barrel cortex would be separated from other cell types by an electrostatic deflection system that diverts droplets into containers based upon their charge. FACS has been used in the central nervous system to study gene expression and function in cerebellar granule cells (Yang et al., 2004) and has successfully isolated a specific subpopulation of neurons from the cerebral cortex for microarray analysis (Arlotta et al., 2005).

7.2 FURTHER WORK

Through these studies 89 genes have been identified that appear to be differentially expressed in the mouse barrel cortex under conditions of vibrissal deprivation that induce neuronal plasticity. This sub-set of genes under microarray analysis show the expected profile of plasticity related genes and it is believed that they should form the basis for future experiments. However QPCR verification of the microarray data has revealed potential false positive results. Therefore additional biological replication of these experiments should be carried out to remove these Type II errors.

None of these genes have been localised to the precise locus of plasticity and CRE-mediated gene expression observed in previous studies (Barth et al., 2000). Therefore a prudent method of validating the microarray data presented here is to carry out *in-situ* hybridisations. Not only would this validate the microarray data but also provide important spatial information regarding the locus of plasticity. Furthermore, the expression of PRGs could act as cellular markers which could be used to dissect out plasticity associated cell types to reduce cellular heterogeneity and the chance of Type I errors.

It may be worth considering analysing the transcriptional response to vibrissal deprivation at later time points. Given that plasticity is maximal at 18 days it is possible that a second wave of gene expression occurs after the 8 day cut-off point used here. It is possible, given the generic function of the majority of the PRGs, observed to be upregulated in the barrel cortex after one day deprivation, that a more plasticity specific set of genes are induced at a later time point. This analysis would be best carried out with microarray.

Following confirmation of differential gene expression, the development of transgenic animals could further understanding of how the PRGs observed here relate to plasticity. For three PRGs identified basigin (*Bsg*), pleiotrophin (*Ptn*) and dystroglycan (*Dg*), transgenic knockout models have already been developed therefore barrel cortex plasticity could be tested in these animals.

As a final point, only approximately 40% of the PRGs identified have been functionally annotated and it is important that these clones are periodically checked, possibly through the NCBI Blast tool to identify whether functional annotation is added.

In concert with a transcriptomic approach for studying differential gene expression in the barrel cortex in plasticity use could be made of another –omic technique. A large number of studies have focused on investigating how individual proteins, biochemical pathways and structural processes alter both the induction and maintenance of synaptic plasticity. However, it is likely that synaptic plasticity involves temporally and spatially coordinated regulation of multiple protein complexes within the activated neural circuit. By using a global proteomics-based approach it is possible that highly diverse protein classes that exhibit altered expression in response vibrissal deprivation conditions that induce plasticity will be revealed. One such approach has already identified changes in the abundance of 79 proteins in the hippocampus under conditions that induce LTP (Mc Nair et al., 2006).

7.3 CONCLUDING REMARKS

It is believed that through the optimisation of microarray technology, conducted in chapter 3, future investigations into gene expression in the mouse barrel cortex using microarray technology, will be free from the constraints of the time taken to optimise this multi-step process. Optimisation of the microarray pipeline has reduced as much error as possible in the data generated here, and for future experiments. Probably the only way to reduce remaining error is through the use of additional biological replication. In

generating a reference RNA sample additional experiments can be easily incorporated into the design used here. This means that future studies investigating gene expression in the mouse barrel cortex can be compared to the experiments conducted here. Whether, these studies involve further time points in the plasticity process, the use of knockout animals at the time points conducted here etc, a practical framework and a bank of expression data regarding the expression levels of genes in the barrel cortex following vibrissal deprivation have been created.

References

Abel T, Nguyen PV, Barad M, Deuel TA, Kandel ER, Bourtchouladze R. (1997) Genetic demonstration of a role for PKA in the late phase of LTP and in hippocampus-based long-term memory. *Cell.* **88**:615-26.

Agmon A, Connors BW. (1991); Thalamocortical responses of mouse somatosensory (barrel) cortex in vitro. *NeuroScience.* **41**:365-79.

Ahissar E, Sosnik R, Haidarliu S. (2000) Transformation from temporal to rate coding in a somatosensory thalamocortical pathway. *Nature.* **406**:302-6.

Ahmed AA, Vias M, Iyer NG, Caldas C, Brenton JD. (2004) Microarray segmentation methods significantly influence data precision. *Nucleic Acids Res.* **32**:e50.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. (2002) Molecular Biology of the Cell. 4th ed., New York: *Garland*

Alfonso J, Pollevick GD, Van Der Hart MG, Flugge G, Fuchs E, Frasch AC.(2004) Identification of genes regulated by chronic psychosocial stress and antidepressant treatment in the hippocampus. *Eur J Neurosci.* **3**:659-66.

Amano M, Ito M, Kimura K, Fukata Y, Chihara K, Nakano T, Matsuura Y, Kaibuchi K. (1996) Phosphorylation and activation of myosin by Rho-associated kinase (Rho-kinase). *J Biol Chem.* **271**:20246-9.

Aoki C, Siekevitz P. (1985) Ontogenetic changes in the cyclic adenosine 3',5'-monophosphate-stimulatable phosphorylation of cat visual cortex proteins, particularly of microtubule-associated protein 2 (MAP 2): effects of normal and dark rearing and of the exposure to light. *J Neurosci.* **5**:2465-83.

Applegate MD, Kerr DS, Landfield PW. (1987) Redistribution of synaptic vesicles during long-term potentiation in the hippocampus. *Brain Res.* **401**:401-6.

Arlotta P, Molyneaux BJ, Chen J, Inoue J, Kominami R, Macklis JD. (2005) Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron*;**45**(2):207-21.

Armstrong-James M, Fox K, Das-Gupta A. (1992) Flow of excitation within rat barrel cortex on striking a single vibrissa. *J Neurophysiol.* **68**:1345-58.

Armstrong-James M, Fox K. (1987) Spatiotemporal convergence and divergence in the rat S1 "barrel" cortex. *J Comp Neurol.* **263**:265-81.

Bailey DJ, Kim JJ, Sun W, Thompson RF, Helmstetter FJ. (1999) Acquisition of fear conditioning in rats requires the synthesis of mRNA in the amygdala. *Behav Neurosci.* **113**:276-82.

Bannister AJ, Kouzarides T. (1996) The CBP co-activator is a histone acetyltransferase. *Nature.* **384**:641-3.

Barco A, Alarcon JM, Kandel ER. (2002) Expression of constitutively active CREB protein facilitates the late phase of long-term potentiation by enhancing synaptic capture. *Cell.* **108**:689-703.

Barrionuevo G, Schottler F, Lynch G. (1980) The effects of repetitive low frequency stimulation on control and "potentiated" synaptic responses in the hippocampus. *Life Sci. 27*:2385-91.

Barth AL; McKenna M, Glazewski S, Hill P, Impey S, Storm D, Fox K. (2000) Upregulation of cAMP response element-mediated gene expression during experience-dependent plasticity in adult neocortex. *JNeurosci.* **20**:4206-16.

Bartsch D, Ghirardi M, Skehel PA, Karl KA, Herder SP, Chen M, Bailey CH, Kandel ER. (1995) Aplysia CREB2 represses long-term facilitation: relief of repression converts transient facilitation into long-term functional and structural change. *Cell.* **83**:979-92.

Battaglia C, Salani G, Consolandi C, Bernardi LR, De Bellis G. (2000) Analysis of DNA microarrays by non-destructive fluorescent staining using SYBR green II. *Biotechniques.* **29**:78-81

Baugh LR, Hill AA, Brown EL, Hunter CP. (2001) Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acids Res.* **29**:E29.

Benjamini Y, Yekutieli D. (2001) The control of the false discovery rate under dependency. *Ann Stat* **29**:1165-1188.

Benson DL, Tanaka H. (1998) N-cadherin redistribution during synaptogenesis in hippocampal Neurons. *J Neurosci.* **18**:6892-904.

Berardi N, Pizzorusso T, Maffei L. (2004) Extracellular matrix and visual cortical plasticity: freeing the synapse. *Neuron.* **44**:905-8.

Bernier G, Pool M, Kilcup M, Alfoldi J, De Repentigny Y, Kothary R. (2000) Acf7 (MACF) is an actin and microtubule linker protein whose expression predominates in neural, muscle, and lung development. *Dev Dyn.* **219**:216-25.

Bisler S, Schleicher A, Gass P, Stehle JH, Zilles K, Staiger JF.(2002) Expression of c-Fos, ICER, Krox-24 and JunB in the whisker-to-barrel pathway of rats: time course of induction upon whisker stimulation by tactile exploration of an enriched environment. *J Chem Neuroanat.* **3**:187-98.

Bito H, Deisseroth K, Tsien RW. (1996) CREB phosphorylation and dephosphorylation: a Ca(2+)- and stimulus duration-dependent switch for hippocampal gene expression. *Cell.* **87**:1203-14.

Black JE, Isaacs KR, Anderson BJ, Alcantara AA, Greenough WT. (1990) Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. *Proc Natl Acad Sci U S A* **87**:5568-72.v

Black, M. A., & Doerge, R. W. (2002). Calculation of the minimum number of replicate spots required for detection of significant gene expression fold change in microarray experiments. *Bioinformatics.* **18**:1609-16.

Blanton JL, D'Ambrozio JA, Sistrunk JE, Midboe EG. (2004) Global changes in the expression patterns of RNA isolated from the hippocampus and cortex of VX exposed mice. *J Biochem Mol Toxicol.* **18**:115-23.

Bliss TV, Lomo T. (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. ***J Physiol.*** **232**:331-56.

Boldrick JC, Alizadeh AA, Diehn M, Dudoit S, Liu CL, Belcher CE, Botstein D, Staudt LM, Brown PO, Relman DA. (2002) Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proc Natl Acad Sci U S A.* **99**:972-7.

Bolliger MF, Frei K, Winterhalter KH, Gloor SM. (2001) Identification of a novel neuroligin in humans which binds to PSD-95 and has a widespread expression. *Biochem J.* **356**:581-8.

Boulton TG, Gregory JS, Cobb MH. (1991) Purification and properties of extra*Cell*ular signal-regulated kinase 1, an insulin-stimulated microtubule-associated protein 2 kinase. *Biochemistry.* **30**:278-86.

Bowtell DD. (1999) Options available--from start to finish--for obtaining expression data by microarray. *Nat Genet.* **21**:25-32.

Brakeman PR, Lanahan AA, O'Brien R, Roche K, Barnes CA, Huganir RL, Worley PF. (1997) Homer: a protein that selectively binds metabotropic glutamate receptors. *Nature.* **386**:284-8.

Brambilla R, Gnesutta N, Minichiello L, White G, Roylance AJ, Herron CE, Ramsey M, Wolfer DP, Cestari V, Rossi-Arnaud C, Grant SG, Chapman PF, Lipp HP, Sturani E, Klein R. (1997) A role for the Ras signalling pathway in synaptic transmission and long-term memory. *Nature.* **390**:281-6.

Brandon EP, Idzerda RL, McKnight GS. (1997) PKA isoforms, neural pathways, and behaviour: making the connection. *Curr Opin Neurobiol.* **7**:397-403.

Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M. (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet.* **29**:365-71.

Brecht M, Preilowski B, Merzenich MM. (1997) Functional architecture of the mystacial vibrissae. Behav *Brain Res.* **84**:81-97.

Brecht M, Sakmann B. (2002) Whisker maps of neuronal subclasses of the rat ventral posterior medial thalamus, identified by whole-Cell voltage recording and morphological reconstruction. *J Physiol.* **538**:495-515.

Browning M, Dunwiddie T, Bennett W, Gispen W, Lynch G. (1979) Synaptic phosphoproteins: specific changes after repetitive stimulation of the hippocampal slice. *Science*. **203**:60-2.

Bruckner G, Brauer K, Hartig W, Wolff JR, Rickmann MJ, Derouiche A, Delpech B, Girard N, Oertel WH, Reichenbach A. (1993) Peri*Neuron*al nets provide a polyanionic, glia-associated form of microenvironment around certain *Neuron*s in many parts of the rat brain. *Glia*. **8**:183-200.

Buckley K, Kelly RB. (1985) Identification of a transmembrane glycoprotein specific for secretory vesicles of neural and endocrine Cells. *J Cell Biol*. **100**:1284-94.

Bustin SA. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol*. **25**:169-93.

Butz S, Okamoto M, Sudhof TC. (1998) A tripartite protein complex with the potential to couple synaptic vesicle exocytosis to *Cell* adhesion in brain. *Cell*. **94**:773-82.

Calakos N, Scheller RH. (1994) Vesicle-associated membrane protein and synaptophysin are associated on the synaptic vesicle. *J Biol Chem*. **269**:24534-7.

Carvell GE, Simons DJ. (1990) Biometric analyses of vibrissal tactile discrimination in the rat. *J Neurosci*. **10**:2638-48

Castellucci VF, Frost WN, Goelet P, Montarolo PG, Schacher S, Morgan JA, Blumenfeld H, Kandel ER. (1986) *Cell* and molecular analysis of long-term sensitization in Aplysia. *J Physiol (Paris)* **81**:349-57.

Cedar H, Schwartz JH. (1972) Cyclic adenosine monophosphate in the nervous system of Aplysia californica. II. Effect of serotonin and dopamine. *J Gen Physiol*. **60**:570-87.

Chang FL, Greenough WT. (1984) Transient and enduring morphological correlates of synaptic activity and efficacy change in the rat hippocampal slice. *Brain Res*. **309**:35-46.

Chawla S, Hardingham GE, Quinn DR, Bading H. CBP: a signal-regulated transcriptional coactivator controlled by nuclear calcium and CaM kinase IV. *Science*. (1998) Sep 4;281(5382):1505-9.

Chen HJ, Rojas-Soto M, Oguni A, Kennedy MB. (1998) A synaptic Ras-GTPase activating protein (p135 SynGAP) inhibited by CaM kinase II. *Neuron.* **20**:895-904

Chen J, Kanai Y, Cowan NJ, Hirokawa N. (1992) Projection domains of MAP2 and tau determine spacings between microtubules in dendrites and axons. *Nature.* **360**:674-7.

Chen X, Zehnbauer B, Gnirke A, Kwok PY. (1997) Fluorescence energy transfer detection as a homogeneous DNA diagnostic method. Proc *Natl Acad Sci U S A.* **94**:10756-61.

Chen Y, Dougherty E R, Bittner M L. (1997) Ratio-based decisions and the quantitative analysis of cDNA microarray images. *Journal of Biomedical Optics*, **2**:364-374.

Cheng G, Mustari MJ, Khanna S, Porter JD. (2003) Comprehensive evaluation of the extraocular muscle critical period by expression profiling in the dark-reared rat and monocularly deprived monkey. *Invest Ophthalmol Vis Sci.* **44**:3842-55.

Cheung VG, Morley M, Aguilar F, Massimi A, Kucherlapati R, Childs G. (1999) Making and reading microarrays. *Nat Genet.* **21**:15-9

Chomczynski P, Sacchi N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem.* **162**:156-9.

Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR, Goodman RH. (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature.* **365**:855-9.

Chung HJ, Xia J, Scannevin RH, Zhang X, Huganir RL. (2000) Phosphorylation of the AMPA receptor subunit GluR2 differentially regulates its interaction with PDZ domain-containing proteins. *J Neurosci.* **20**:7258-67

Churchill GA. (2002) Fundamentals of experimental design for cDNA microarrays. *Nat Genet.* **32**:490-5.

Cleveland, W.S. (1979) Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, Vol. 74, pp. **829**-836.

Cleveland, W.S. and Devlin, S.J. (1988) Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, Vol. 83, pp. **596**-610.

Coan EJ, Saywood W, Collingridge GL. (1987) MK-801 blocks NMDA receptor-mediated synaptic transmission and long term potentiation in rat hippocampal slices. *Neurosci Lett.* **80**:111-4.

Cobb MH, Robbins DJ, Boulton TG. (1991) ERKs, extra*Cell*ular signal-regulated MAP-2 kinases. *Curr Opin Cell Biol.* **3**:1025-32.

Cochran and G. M. Cox. (1957). Experimental Design. John Wiley & Sons, *New York, NY.* **p397**

Colantuoni C, PurCell AE, Bouton CM, Pevsner J. (2000) High throughput analysis of gene expression in the human brain. *J Neurosci* Res. **59**:1-10.

Collingridge GL, Kehl SJ, McLennan H. (1983) The antagonism of amino acid-induced excitations of rat hippocampal CA1 neurones in vitro. *J Physiol.* **334**:19-31.

Conn PJ, and Pin JP. (1997). Pharmacology and functions of metabotropic glutamate receptors. *Annual Review of Pharmacology and Toxicology* **37**: 207-237

Costigan M, Befort K, Karchewski L, Griffin RS, D'Urso D, Allchorne A, Sitarski J, Mannion JW, Pratt RE, Woolf CJ. (2002) Replicate high-density rat genome oligonucleotide microarrays reveal hundreds of regulated genes in the dorsal root ganglion after peripheral nerve injury. *BMC Neuroscience.* **3**:16.

Crino PB, Eberwine J. (1996) Molecular characterization of the dendritic growth cone: regulated mRNA transport and local protein synthesis. *Neuron.* **17**:1173-87.

Cruzalegui FH, Kapiloff MS, Morfin JP, Kemp BE, Rosenfeld MG, Means AR. (1992) Regulation of intrasteric inhibition of the multifunctional calcium/calmodulin-dependent protein kinase. *Proc Natl Acad Sci U S A.* **89**:12127-31.

Dai Z, Peng HB. (1998) Fluorescence microscopy of calcium and synaptic vesicle dynamics during synapse formation in tissue culture. *Histochem J.* **3**:189-96.

D'Angelo E, Rossi P, Gall D, Prestori F, Nieus T, Maffei A, Sola E. (2004); Long-term potentiation of synaptic transmission at the mossy fiber-granule Cell relay of cerebellum. *Prog Brain Res.* **148**:69-80.

Dash PK, Hochner B, Kandel ER. (1990) Injection of the cAMP-responsive element into the nucleus of Aplysia sensory neurons blocks long-term facilitation. *Nature* **345**(6277):718-21.

Dash PK, Karl KA, Colicos MA, Prywes R, Kandel ER. (1991) cAMP response element-binding protein is activated by Ca2+/calmodulin- as well as cAMP-dependent protein kinase. *Proc Natl Acad Sci U S A.* **88**:5061-5.

Deegan, O. Bakajin, T. F. Dupont, G. Huber, S. R. Nagel, and T. A. Witten, "Capillary flow as the cause of ring stains from dried liquid drops" *Nature* 389, 6653 (1997).

Deisseroth K, Bito H, Tsien RW. (1996) Signaling from synapse to nucleus: postsynaptic CREB phosphorylation during multiple forms of hippocampal synaptic plasticity. *Neuron.* **16**:89-101.

Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **4**:P3.

DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM. (1996) Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet.* **14**:457-60.

DeRisi JL, Iyer VR, Brown PO. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science.* **278**:680-6.

Desai A, Mitchison TJ. (1997) Microtubule polymerization dynamics. *Annu Rev Cell Dev Biol.* **13**:83-117.

Desmond NL, Levy WB. (1986) Changes in the postsynaptic density with long-term potentiation in the dentate gyrus. *J Comp Neurol* **253**:476-82.

Desmond NL, Levy WB. (1990) Morphological correlates of long-term potentiation imply the modification of existing synapses, not synaptogenesis, in the hippocampal dentate gyrus. *Synapse.* **5**:139-43.

Devoogd TJ, Nixdorf B, Nottebohm F. (1985) Synaptogenesis and changes in synaptic morphology related to acquisition of a new behavior. *Brain Res.* **329**:304-8.

D'haeseleer P, Liang S, Somogyi R. (2000) Genetic network inference: from co-expression clustering to reverse engineering. *Bioinformatics.* 16:707-26.

Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A. (2004) Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques.* **37**:112-4, 116, 118-9.

Diamond ME, Armstrong-James M, Ebner FF. (1992) Somatic sensory responses in the rostral sector of the posterior group (POm) and in the ventral posterior medial nucleus (VPM) of the rat thalamus. *J Comp Neurol.* **318**:462-76.

Diatchenko L, Lau YF, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, Siebert PD. (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci U S A.* **93**:6025-30.

Diaz-Nido J, Montoro RJ, Lopez-Barneo J, Avila J. (1993) High external potassium induces an increase in the phosphorylation of the cytoskeletal protein MAP2 in rat hippocampal slices. Eur *J Neurosci.* **5**:818-24.

Dickson BJ. (2001) Rho GTPases in growth cone guidance. *Curr Opin Neurobiol.* **11**:103-10.

Dingledine R, Borges K, Bowie D, Traynelis SF. (1999) The glutamate receptor ion channels. *Pharmacol Rev.* **51**:7-61.

Dityatev AE, Bolshakov VY. (2005) Amygdala, Long-term Potentiation, and Fear Conditioning. *Neuroscientist.* **11**:75-88.

Dityatev A, Schachner M. (2003) Extracellular matrix molecules and synaptic plasticity. *Nat Rev Neuroscience* **4**:456-68.

Dobbin K, Shih JH, Simon R. (2003) Statistical design of reverse dye microarrays. *Bioinformatics.* **19**:803-10.

Dobbin K, Simon R. (2002) Comparison of microarray designs for class comparison and class discovery. *Bioinformatics.* **18**:1438-45.

Donoghue JP and Wise SP. (1982) The motor cortex of the rat: cytoarchitecture and microstimulation mapping. *J Comp Neurol* **212**: 76–88.

Dorfl J. (1982) The musculature of the mystacial vibrissae of the white mouse. *J Anat.* **135**:147-54

Draghici S, Kulaeva O, Hoff B, Petrov A, Shams S, Tainsky MA. (2003) Noise sampling method: an ANOVA approach allowing robust selection of differentially regulated genes measured by DNA microarrays. *Bioinformatics.* **19**:1348-59.

Draghici, S. (2003). *Data Analysis Tools for DNA Microarrays*. Chapman and Hall, London.

Drewes G, Ebneth A, Mandelkow EM. (1998) MAPs, MARKs and microtubule dynamics. Trends Biochem Sci. **23**:307-11.

Dubnau J, Tully T. (1998) Gene discovery in Drosophila: new insights for learning and memory. *Annu Rev Neurosci.* **21**:407-44.

Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM. (1999) Expression profiling using cDNA microarrays. *Nat Genet.* **21**:10-4.

Durand GM, Kovalchuk Y, Konnerth A. (1996) Long-term potentiation and functional synapse induction in developing hippocampus. *Nature.* **381**(6577):71-5.

Eberwine J, Yeh H, Miyashiro K, Cao Y, Nair S, Finnell R, Zettel M, Coleman P. (1992) Analysis of gene expression in single live neurons. *Proc Natl Acad Sci U S A.* **89**:3010-4.
Eberwine J, Yeh H, Miyashiro K, Cao Y, Nair S, Finnell R, Zettel M, Coleman P. (1992) Analysis of gene expression in single live neurons. *Proc Natl Acad Sci U S A.* **89**:3010-4.

Eisen MB, Spellman PT, Brown PO, Botstein D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A.* **95**:14863-8.

Eisen, M.B. and Brown, P.O. (1999). DNA arrays for analysis of gene expression. *Methods Enzymol.* **303**:179-205

Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, Weiss RA, Liotta LA. (1996) Laser capture microdissection. *Science.* **274**(5289):998-1001.

Engert F, Bonhoeffer T. (1999) Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature.* May **399**:66-70.

English JD, Sweatt JD. (1997) A requirement for the mitogen-activated protein kinase cascade in hippocampal long term potentiation. *J Biol Chem.* **272**:19103-6.

Enslen H, Sun P, Brickey D, Soderling SH, Klamo E, Soderling TR. (1994) Characterization of Ca2+/calmodulin-dependent protein kinase IV. Role in transcriptional regulation. *J Biol Chem.* **269**:15520-7.

Ernst LA, Gupta RK, Mujumdar RB, Waggoner AS. (1989) Cyanine dye labeling reagents for sulfhydryl groups. *Cytometry.* **10**:3-10.

Erzurumlu RS, Jhaveri S. (1990) Thalamic axons confer a blueprint of the sensory periphery onto the developing rat somatosensory cortex. *Brain Res* Dev *Brain Res.* **56**:229-34.

Feldmeyer D, Egger V, Lubke J, Sakmann B. (1999) Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single 'barrel' of developing rat somatosensory cortex. *J Physiol.* **521** Pt 1:169-90.

Feldmeyer D, Lubke J, Silver RA, Sakmann B. (2002) Synaptic connections between layer 4 spiny neurone-layer 2/3 pyramidal Cell pairs in juvenile rat barrel cortex: physiology and anatomy of interlaminar signalling within a cortical column. *J Physiol.* **538**:803-22.

Felinski EA, Kim J, Lu J, Quinn PG. (2001) Recruitment of an RNA polymerase II complex is mediated by the constitutive activation domain in CREB, independently of CREB phosphorylation. *Mol Cell Biol.* **21**:1001-10.

Felinski EA, Quinn PG. (1999) The CREB constitutive activation domain interacts with TATA-binding protein-associated factor 110 (TAF110) through specific hydrophobic residues in one of the three subdomains required for both activation and TAF110 binding. *J Biol Chem.* **274**:11672-8

Feng Y, Liang HL, Wong-Riley M (2004) Differential gene expressions in the visual cortex of postnatal day 1 versus day 21 rats revealed by suppression subtractive hybridization. *Gene.*;**329**:93-101.

Ferreira A, Chin LS, Li L, Lanier LM, Kosik KS, Greengard P. (1998) Distinct roles of synapsin I and synapsin II during *Neurona*l development. *Mol Med.* **4**:22-8.

Finkbeiner S, Tavazoie SF, Maloratsky A, Jacobs KM, Harris KM, Greenberg ME. (1997) CREB: a major mediator of neuronal neurotrophin responses. *Neuron.* **19**:1031-47.

Fisher, R. A. (1951) in The Design of Experiments 6th Ed. *Oliver and Boyd, Edinburgh*

Fox K. (1992) A critical period for experience-dependent synaptic plasticity in rat barrel cortex. *JNeurosci.* **12**:1826-38.

Fox K. (1995) The critical period for long-term potentiation in primary sensory cortex. *Neuron.* **15**:485-8.

Fox K (2002) Anatomical pathways and molecular mechanisms for plasticity in the barrel cortex. *Neuroscience.* **111**(4):799-814.

Frangakis MV, Ohmstede CA, Sahyoun N. (1991) A brain-specific Ca2+/calmodulin-dependent protein kinase (CaM kinase-Gr) is regulated by autophosphorylation. Relevance to neuronal Ca2+ signaling. *J Biol Chem.* **266**:11309-16.

Frey U, Frey S, Schollmeier F, Krug M. (1996) Influence of actinomycin D, a RNA synthesis inhibitor, on long-term potentiation in rat hippocampal *Neuron*s in vivo and in vitro. *J Physiol.* **490**:703-11.

Frey U, Krug M, Reymann KG, and Matthies H. (1988) Anisomycin, an inhibitor of protein synthesis, blocks late phases of LTP phenomena in the hippocampal CA1 region in vitro. *Brain Res* **452**: 57-65,

Frey U, Morris RG. (1998) Weak before strong: dissociating synaptic tagging and plasticity-factor accounts of late-LTP. *Neuropharmacology.* **37**:545-52.

Frodin M, Gammeltoft S. (1999) Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Mol Cell Endocrinol.* **151**:65-77.

Fujii S, Saito K, Miyakawa H, Ito K, Kato H. (1991) Reversal of long-term potentiation (depotentiation) induced by tetanus stimulation of the input to CA1 neurons of guinea pig hippocampal slices. *Brain Res.* **555**:112-22.

Gamblin TC, Nachmanoff K, Halpain S, Williams RC Jr. (1996) Recombinant microtubule-associated protein 2c reduces the dynamic instability of individual microtubules. *Biochemistry.* **35**:12576-86.

Gao P, Bermejo R, and Zeigler HP. (2001). Vibrissa deaffentation and rodent whisking patterns: behavioral evidence for a central pattern generator. *J Neurosci* **21**: 5374–5380

Garner CC, Zhai RG, Gundelfinger ED, Ziv NE. (2002) Molecular mechanisms of CNS synaptogenesis. *Trends Neurosci.* **25**:243-51.

Geppert M, Goda Y, Hammer RE, Li C, Rosahl TW, Stevens CF, Sudhof TC. (1994) Synaptotagmin I: a major Ca2+ sensor for transmitter release at a central synapse. *Cell.* **79**:717-27.

Gibbs PJ, Cameron C, Tan LC, Sadek SA, Howell WM. (2003) House keeping genes and gene expression analysis in transplant recipients: a note of caution. *Transpl Immunol.* **12**:89-97

Gibson JM, Welker WI. (1983); Quantitative studies of stimulus coding in first-order vibrissa afferents of rats. 1. Receptive field properties and threshold distributions. *Somatosens Res.* **1**:51-67.

Gibson JM, Welker WI. (1983); Quantitative studies of stimulus coding in first-order vibrissa afferents of rats. 2. Adaptation and coding of stimulus parameters. *Somatosens Res.* **1**:95-117.

Glazewski S, Barth AL, Wallace H, McKenna M, Silva A, Fox K. (1999) Impaired experience-dependent plasticity in barrel cortex of mice lacking the alpha and delta isoforms of CREB. *Cereb Cortex* . **9**:249-56.

Glazewski S, Chen CM, Silva A, Fox K. (1996) Requirement for alpha-CaMKII in experience-dependent plasticity of the barrel cortex. *Science.* **272**:421-3.

Glazewski S, McKenna M, Jacquin M, Fox K. (1998) Experience-dependent depression of vibrissae responses in adolescent rat barrel cortex. Eur *J Neurosci.* **10**:2107-16.

Glazewski S., Fox K. (1996) The time-course of experience-dependent synaptic potentiation and depression in barrel cortex of adolescent rats. *J. Neurophysiol.*, **7**: 1714-1730.

Goldberg Y. (1999) Protein phosphatase 2A: who shall regulate the regulator? *Biochem Pharmacol.* **57**:321-8.

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science.* **286**:531-7.

Gonzalez GA, Montminy MR. (1989) Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell.* **59**:675-80

Gorecki DC, Lukasiuk K, Szklarczyk A, Kaczmarek L. (1998) Kainate-evoked changes in dystrophin messenger RNA levels in the rat hippocampus. *Neuroscience.* **84**(2):467-77.

Greengard, P., F. Benfenati, and F. Valtorta. (1994). Synapsin I, an actin-binding protein regulating synaptic vesicle traffic in the nerve terminal.

Gris P, Murphy S, Jacob JE, Atkinson I, Brown A. (2003) Differential gene expression profiles in embryonic, adult-injured and adult-uninjured rat spinal cords. *Mol Cell Neurosci.* **24:**555-67.

Guerini D. (1997) Calcineurin: not just a simple protein phosphatase. *Biochem Biophys Res Commun.* **235**:271-5.

Gundersen GG, Cook TA. (1999) Microtubules and signal transduction. *Curr Opin Cell Biol.* **11**:81-94.

Gustafsson B, Wigstrom H, Abraham WC, Huang YY. (1987) Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. *J Neurosci.* **7**:774-80.

Guzowski JF, Lyford GL, Stevenson GD, Houston FP, McGaugh JL, Worley PF, Barnes CA. (2000) Inhibition of activity-dependent arc protein expression in the rat hippocampus impairs the maintenance of long-term potentiation and the consolidation of long-term memory. *J Neurosci.* **20**:3993-4001.

Hacia JG. (1999) Resequencing and mutational analysis using oligonucleotide microarrays. *Nat Genet.* **21**:42-7.

Hand PJ (1982) Plasticity of the rat cortical barrel system. In: Changing concepts of the nervous system (Morrison AR, Strick PL, eds), pp 49-68. New York: Academic.

Hata Y, Butz S, Sudhof TC. (1996) CASK: a novel dlg/PSD95 homolog with an N-terminal calmodulin-dependent protein kinase domain identified by interaction with neurexins. *J Neurosci.* **16**:2488-94.

Hay JC. (2001) SNARE complex structure and function. *Exp Cell Res.* **271**:10-21.

Hayashi Y, Shi SH, Esteban JA, Piccini A, Poncer JC, Malinow R. (2000) Driving AMPA receptors into synapses by LTP and CaMKII: requirement for GluR1 and PDZ domain interaction. *Science.* **287**:2262-7.

Hebb, D.O. (1949) The organization of behavior. A neuropsychological theory. *New York: John Wiley.*

Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N, Quackenbush J. (2000) A concise guide to cDNA microarray analysis. *Biotechniques.* **29**:548-50, 552-4, 556 passim.

Herzenberg LA, Sweet RG, Herzenberg LA. (1976) Fluorescence-activated cell sorting. *Sci Am.* **234**(3):108-17.

Hinds HL, Tonegawa S, Malinow R. (1998) CA1 long-term potentiation is diminished but present in hippocampal slices from alpha-CaMKII mutant mice. *Learn Mem.* **5**:344-54.

Hiyoshi M, Hosoi S. (1994) Assay of DNA denaturation by polymerase chain reaction-driven fluorescent label incorporation and fluorescence resonance energy transfer. *Anal Biochem.* **221**:306-11.

Hoen PA, de Kort F, van Ommen GJ, den Dunnen JT. (2003) Fluorescent labelling of cRNA for microarray applications. *Nucleic Acids Res.* **3**:e20.

Holland PM, Abramson RD, Watson R, Gelfand DH. (1991) Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc Natl Acad Sci U S A.* **88**:7276-80.

Hollmann M, Heinemann S. (1994); Cloned glutamate receptors. *Annu Rev Neurosci.* **17**:31-108.

Holloway AJ, van Laar RK, Tothill RW, Bowtell DD. (2002) Options available--from start to finish--for obtaining data from DNA microarrays II. *Nat Genet.* **32**:481-9.

Hong SJ, Li H, Becker KG, Dawson VL, Dawson TM. (2004) Identification and analysis of plasticity-induced late-response genes. *Proc Natl Acad Sci U S A.* **101**:2145-50.

Hosack DA, Dennis G Jr, Sherman BT, Lane HC, Lempicki RA. (2003); Identifying biological themes within lists of genes with EASE. *Genome Biol.* **4**:R70.

Hosaka M, Hammer RE, Sudhof TC. (1999) A phospho-switch controls the dynamic association of synapsins with synaptic vesicles. *Neuron.* **24**:377-87.

Hubel DH, Wiesel TN. (1970) The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *J Physiol.* **206**:419-36.

Hubel DH, Wiesel TN, LeVay S. (1977) Plasticity of ocular dominance columns in monkey striate cortex. *Philos Trans R Soc Lond B Biol Sci.* **278**:377-409.

Hughes PE, Alexi T, Walton M, Williams CE, Dragunow M, Clark RG, Gluckman PD (1999) Activity and injury-dependent expression of inducible transcription factors, growth factors and apoptosis-related genes within the central nervous system. *Prog Neurobiol.* **57**:421-50.

Hummler E, Cole TJ, Blendy JA, Ganss R, Aguzzi A, Schmid W, Beermann F, Schutz G. (1994) Targeted mutation of the CREB gene: compensation within the CREB/ATF family of transcription factors. *Proc Natl Acad Sci U S A.* **91**:5647-51.

Huntley GW, Benson DL. (1999) Neural (N)-cadherin at developing thalamocortical synapses provides an adhesion mechanism for the formation of somatopically organized connections. *J Comp Neurol.* **407**:453-71.

Impey S, Mark M, Villacres EC, Poser S, Chavkin C, Storm DR. (1996) Induction of CRE-mediated gene expression by stimuli that generate long-lasting LTP in area CA1 of the hippocampus. *Neuron.* **16**:973-82.

Impey S, Obrietan K, Storm DR. (1999) Making new connections: role of ERK/MAP kinase signaling in neuronal plasticity. *Neuron.* **23**:11-4.

Impey S, Obrietan K, Wong ST, Poser S, Yano S, Wayman G, Deloulme JC, Chan G, Storm DR. (1998) Cross talk between ERK and PKA is required for Ca2+ stimulation of CREB-dependent transcription and ERK nuclear translocation. *Neuron.* **21**:869-83.

Inokuchi K, Kato A, Hiraia K, Hishinuma F, Inoue M, Ozawa F. (1996) Increase in activin beta A mRNA in rat hippocampus during long-term potentiation. *FEBS Lett.* **382**:48-52.

Inoue A, Sanes JR. (1997) Lamina-specific connectivity in the brain: regulation by N-cadherin, neurotrophins, and glycoconjugates. *Science.* **276**:1428-31.

Irie M, Hata Y, Takeuchi M, Ichtchenko K, Toyoda A, Hirao K, Takai Y, Rosahl TW, Sudhof TC. (1997) Binding of neuroligins to PSD-95. *Science.* **277**:1511-5.

Isaac JT, Nicoll RA, Malenka RC. (1995) Evidence for silent synapses: implications for the expression of LTP. *Neuron* **15**:427-34.

Ishii, T., K. Moriyoshi, H. Sugihara, K. Sakurada, H. Kaditani, M. Yokoi, C. Akazawa, R. Shigemoto, N. Mizuno, M. Masu. (1993). Molecular characterization of the family of the N-methyl-D-aspartate receptor subunits. *Journal of Biological Chemistry.* **268**: 2836-2843.

Jacquin MF, Renehan WE, Rhoades RW, Panneton WM. (1993) Morphology and topography of identified primary afferents in trigeminal subnuclei principalis and oralis. *J Neurophysiol.* **70**:1911-36.

Jacquin MF, Rhoades RW. (1990) Cell structure and response properties in the trigeminal subnucleus oralis. *Somatosens Mot Res.* **7**:265-88.

Jahn R, Lang T, Sudhof TC. (2003) Membrane fusion. *Cell.* **112**:519-33.

Jahn R, Sudhof TC. (1999); Membrane fusion and exocytosis. *Annu Rev Biochem.* **68**:863-911.

Jain AN, Tokuyasu TA, Snijders AM, Segraves R, Albertson DG, Pinkel D. (2002) Fully automatic quantification of microarray image data. *Genome Res.* **12**:325-32.

Jenson SD, Robetorye RS, Bohling SD, Schumacher JA, Morgan JW, Lim MS, Elenitoba-Johnson KS. (2003) Validation of cDNA microarray gene expression data obtained from linearly amplified RNA. *Mol Pathol.* **56**:307-12.

Jian B, Jones PL, Li Q, Mohler ER 3rd, Schoen FJ, Levy RJ. (2001) Matrix metalloproteinase-2 is associated with tenascin-C in calcific aortic stenosis. *Am J Pathol.* **159**:321-7.

Jin W, Riley RM, Wolfinger RD, White KP, Passador-Gurgel G, Gibson G. (2001) The contributions of sex, genotype and age to transcriptional variance in Drosophila melanogaster. *Nat Genet.* **29**:389-95.

Kaibuchi K, Kuroda S, Amano M. (1999); Regulation of the cytoskeleton and *Cell* adhesion by the Rho family GTPases in mammalian Cells. *Annu Rev Biochem.* **68**:459-86.

Kamme F, Salunga R, Yu J, Tran DT, Zhu J, Luo L, Bittner A, Guo HQ, Miller N, Wan J, Erlander M. (2003) Single-cell microarray analysis in hippocampus CA1: demonstration and validation of cellular heterogeneity. *JNeurosci.* **23**:3607-15.

Kandel ER, Schwartz JH. (1982) Molecular biology of learning: modulation of transmitter release. *Science.* **218**:433-43.

Kanehisa M, Goto S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**:27-30.

Kanekura T, Miyauchi T, Tashiro M, Muramatsu T. (1991) Basigin, a new member of the immunoglobulin superfamily: genes in different mammalian species, glycosylation changes in the molecule from adult organs and possible variation in the N-terminal sequences. *Cell Struct Funct.* **16**:23-30.

Kataoka H, DeCastro R, Zucker S, Biswas C. (1993) Tumor cell-derived collagenase-stimulatory factor increases expression of interstitial collagenase, stromelysin, and 72-kDa gelatinase. *Cancer Res.* **53**:3154-8.

Katz B. (1969). The Release of Neural Transmitter Substances. Springfield, IL: C C Thomas.

Kawasaki H, Springett GM, Mochizuki N, Toki S, Nakaya M, Matsuda M, Housman DE, Graybiel AM. (1998) A family of cAMP-binding proteins that directly activate Rap1. *Science.* **282**:2275-9.

Kelso SR, Ganong AH, Brown TH. (1986) Hebbian synapses in hippocampus. *Proc Natl Acad Sci U S A.* **83**:5326-30.

Kerr MK, Churchill GA. (2001) Bootstrapping cluster analysis: assessing the reliability of conclusions from microarray experiments. *Proc Natl Acad Sci U S A.* **98**:8961-5.

Kerr MK, Churchill GA. (2001) Experimental design for gene expression microarrays. *Biostatistics.* **2**:183-201.

Kerr MK, Churchill GA. (2001) Statistical design and the analysis of gene expression microarray data. *Genet Res.* **77**:123-8.

Kerr MK, Martin M, Churchill GA. (2000) Analysis of variance for gene expression microarray data. *J Comput Biol.* **7**:819-37

Kim JH, Liao D, Lau LF, Huganir RL. (1998) SynGAP: a synaptic RasGAP that associates with the PSD-95/SAP90 protein family. *Neuron.* **20**:683-91.

Kinnunen T, Raulo E, Nolo R, Maccarana M, Lindahl U, Rauvala H. (1996) Neurite outgrowth in brain *Neuron*s induced by heparin-binding growth-associated molecule (HB-GAM) depends on the specific interaction of HB-GAM with heparan sulfate at the cell surface. *J Biol Chem.* **271**:2243-8.

Kleim JA, Lussnig E, Schwarz ER, Comery TA, Greenough WT. (1996) Synaptogenesis and Fos expression in the motor cortex of the adult rat after motor skill learning. *J Neurosci.* **16**:4529-35.
Kleim JA, Vij K, Ballard DH, Greenough WT. (1997) Learning-dependent synaptic modifications in the cerebellar cortex of the adult rat persist for at least four weeks. *J Neurosci.* **17**:717-21.

Klein M, Kandel ER. (1980) Mechanism of calcium current modulation underlying presynaptic facilitation and behavioral sensitization in Aplysia. *Proc Natl Acad Sci USA.* **77**:6912-6.

Ko MS, Kitchen JR, Wang X, Threat TA, Wang X, Hasegawa A, Sun T, Grahovac MJ, Kargul GJ, Lim MK, Cui Y, Sano Y, Tanaka T, Liang Y, Mason S, Paonessa PD, Sauls AD, DePalma GE, Sharara R, Rowe LB, Eppig J, Morrell C, Doi H. (2000) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development. *Development.* **127**:1737-49.

Koppe G, Bruckner G, Hartig W, Delpech B, Bigl V. (1997) Characterization of proteoglycan-containing peri*Neuron*al nets by enzymatic treatments of rat brain sections. *Histochem J.* **29**:11-20.

Kornau HC, Seeburg PH, Kennedy MB. (1997) Interaction of ion channels and receptors with PDZ domain proteins. *Curr Opin Neurobiol.* 7:368-73.

Kornau HC, Seeburg PH. (1997) Partner selection by PDZ domains. *Nat Biotechnol.* **15**:319.

Kornhauser JM, Greenberg ME. (1997) A kinase to remember: dual roles for MAP kinase in long-term memory. *Neuron.* **18**:839-42.

Kowalski RJ, Williams RC Jr. (1993) Microtubule-associated protein 2 alters the dynamic properties of microtubule assembly and disassembly. *J Biol Chem.* **268**:9847-55.

Kozma R, Ahmed S, Best A, Lim L. (1995) The Ras-related protein Cdc42Hs and bradykinin promote formation of peripheral actin microspikes and filopodia in Swiss 3T3 fibroblasts. *Mol Cell Biol.* **15**:1942-52.

Krug M, Lossner B, Ott T. (1984) Anisomycin blocks the late phase of long-term potentiation in the dentate gyrus of freely moving rats. *Brain Res Bull.* **13**:39-42.

Kwok RP, Lundblad JR, Chrivia JC, Richards JP, Bachinger HP, Brennan RG, Roberts SG, Green MR, Goodman RH. (1994) Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature.* **370**:223-6.

Lachance PE, Chaudhuri A. (2004) Microarray analysis of developmental plasticity in monkey primary visual cortex. *J Neurochem.* **88**:1455-69.

Lanahan, A. A. & Worley, P. F. (1998) Immediate-early genes and synaptic function.

Landegent JE, Jansen in de Wal N, Dirks RW, Baao F, van der Ploeg M. (1987) Use of whole cosmid cloned genomic sequences for chromosomal localization by non-radioactive in situ hybridization. *Hum Genet.* **77**:366-70.

Lauri SE, Rauvala H, Kaila K, Taira T. (1998) Effect of heparin-binding growth-associated molecule (HB-GAM) on synaptic transmission and early LTP in rat hippocampal slices. *Eur J Neuroscience.* **10**:188-94.

LeClerc N, Kosik KS, Cowan N, Pienkowski TP, Baas PW. (1993) Process formation in Sf9 *Cells* induced by the expression of a microtubule-associated protein 2C-like construct. *Proc Natl Acad Sci U S A.* **90**:6223-7.

Lee CH, Herman T, Clandinin TR, Lee R, Zipursky SL. (2001) N-cadherin regulates target specificity in the Drosophila visual system. *Neuron.* **30**:437-50.

Lee HK, Takamiya K, Han JS, Man H, Kim CH, Rumbaugh G, Yu S, Ding L, He C, Petralia RS, Wenthold RJ, Gallagher M, Huganir RL. (2003) Phosphorylation of the AMPA receptor GluR1 subunit is required for synaptic plasticity and retention of spatial memory. *Cell.* **112**:631-43.

Lee LG, Connell CR, Bloch W. (1987); Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**:335-50.

Lee ML, Kuo FC, Whitmore GA, Sklar J. (2000) Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proc Natl Acad Sci U S A.* **97**:9834-9.

Lengauer C, Riethman H, Cremer T. (1990) Painting of human chromosomes with probes generated from hybrid cell lines by PCR with Alu and L1 primers. *Hum Genet.* **86**:1-6.

Leung YF, Cavalieri D. (2003) Fundamentals of cDNA microarray data analysis. *Trends Genet.* **19**:649-59.

Levenson JM, Weeber EJ, Sweatt JD, Eskin A. (2002) Glutamate uptake in synaptic plasticity: from mollusc to mammal. *Curr Mol Med.* **2**:593-603.

Levy CC. (1975) Roles of RNases in cellular regulatory mechanisms. *Life Sci.* **17**:311-6.

Levy WB, Steward O. (1979) Synapses as associative memory elements in the hippocampal formation. *Brain Res.* **175**:233-45.

Li P, Kerchner GA, Sala C, Wei F, Huettner JE, Sheng M, Zhuo M. (1999) AMPA receptor-PDZ interactions in facilitation of spinal sensory synapses. *Nat Neurosci.* **2**:972-7.

Liang P, Pardee AB. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science.* **257**:967-71.

Liao D, Hessler NA, Malinow R. (1995) Activation of postsynaptically silent synapses during pairing-induced LTP in CA1 region of hippocampal slice. *Nature.* **375**:400-4.

Livak KJ, Flood SJ, Marmaro J, Giusti W, Deetz K. (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl.* **4**:357-62.

Livesey FJ, Hunt SP. (1996) Identifying changes in gene expression in the nervous system: mRNA differential display. *Trends Neurosci.* **19**:84-8.

Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol.* **14**:1675-80

Lomo T (1966) Frequency potentiation of excitatory synaptic activity in the dentate area of the hippocampal formation. *Acta Physiol Scand* **68**:128.

Long AD, Mangalam HJ, Chan BY, Tolleri L, Hatfield GW, Baldi P. (2001) Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in Escherichia coli K12. *J Biol Chem.* **276**:7-44.
Lonze BE, Ginty DD. (2002) Function and regulation of CREB family transcription factors in the nervous system. *Neuron* **35**:605-23.

Luo L, Salunga RC, Guo H, Bittner A, Joy KC, Galindo JE, Xiao H, Rogers KE, Wan JS, Jackson MR, Erlander MG.(1999) Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat Med.* **1**:117-22.

Luo L. (2000) Rho GTPases in Neuronal morphogenesis. *Nat Rev Neurosci.* **1**:173-80.

Luo L. (2002) Actin cytoskeleton regulation in *Neuron*al morphogenesis and structural plasticity. *Annu Rev Cell Dev Biol.* **18**:601-35.

Luo Z, Geschwind DH. (2001) Microarray applications in neuroscience. *Neurobiol Dis.* **8**:183-93.

Lyford GL, Yamagata K, Kaufmann WE, Barnes CA, Sanders LK, Copeland NG, Gilbert DJ, Jenkins NA, Lanahan AA, Worley PF. (1995) Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. *Neuron.* **14**:433-45.

Lynch MA. (2004) Long-term potentiation and memory. *Physiol Rev.* **84**:87-136.

Ma PM. (1991) The barrelettes—architectonic vibrissal representations in the brainstem trigeminal complex of the mouse. I. Normal structural organization. *J Comp Neurol* **309**: 161-199.

Maeda N, Nishiwaki T, Shintani T, Hamanaka H, Noda M. (1996) 6B4 proteoglycan/phosphacan, an extracellular variant of receptor-like protein-tyrosine phosphatase zeta/RPTPbeta, binds pleiotrophin/heparin-binding growth-associated molecule (HB-GAM). *J Biol Chem.* **271**:21446-52.

Mandelkow EM, Biernat J, Drewes G, Steiner B, Lichtenberg-Kraag B, Wille H, Gustke N, Mandelkow E. (1993) Microtubule-associated protein tau, paired helical filaments, and phosphorylation. *Ann N Y Acad Sci.* **695**:209-16.

Martin KC, Casadio A, Zhu H, Yaping E, Rose JC, Chen M, Bailey CH, Kandel ER. (1997) Synapse-specific, long-term facilitation of aplysia sensory to motor synapses: a function for local protein synthesis in memory storage. *Cell.* **91**:927-38.

Masayesva BG, Ha P, Garrett-Mayer E, Pilkington T, Mao R, Pevsner J, Speed T, Benoit N, Moon CS, Sidransky D, Westra WH, Califano J. (2004) Gene expression alterations over large chromosomal regions in cancers include multiple genes unrelated to malignant progression. *Proc Natl Acad Sci U S A.* **101**:8715-20.

Mataga N, Fujishima S, Condie BG, Hensch TK. (2001) Experience-dependent plasticity of mouse visual cortex in the absence of the neuronal activity-dependent marker egr1/zif268. *JNeurosci.* **21**:9724-32.

Matsuda S, Mikawa S, Hirai H. (1999) Phosphorylation of serine-880 in GluR2 by protein kinase C prevents its C terminus from binding with glutamate receptor-interacting protein. *J Neurochem.* **73**:1765-8.

Matsuo R, Kato A, Sakaki Y, Inokuchi K. (1998) Cataloging altered gene expression during rat hippocampal long-term potentiation by means of differential display. *Neurosci Lett.* **244**:173-6.

Matthews RP, Guthrie CR, Wailes LM, Zhao X, Means AR, McKnight GS. (1994) Calcium/calmodulin-dependent protein kinase types II and IV differentially regulate CREB-dependent gene expression. *Mol Cell Biol.* **14**:6107-16.

Maximov A, Sudhof TC, Bezprozvanny I. (1999) Association of Neuronal calcium channels with modular adaptor proteins. *J Biol Chem.* **274**:24453-6

Maycox PR, Deckwerth T, Hell JW, Jahn R. . (1988) Glutamate uptake by brain synaptic vesicles. Energy dependence of transport and functional reconstitution in proteoliposomes. *J Biol Chem* **263**:15423-8.

Mayer ML, Westbrook GL. (1987) Permeation and block of N-methyl-D-aspartic acid receptor channels by divalent cations in mouse cultured central neurones. *J Physiol.* **394**:501-27.

Mayer P, Ammon S, Braun H, Tischmeyer H, Riechert U, Kahl E, Hollt V.(2002) Gene expression profile after intense second messenger activation in cortical primary neurones. *J Neurochem.* **82**(5):1077-86.

Mayr B, Montminy M. (2001) Transcriptional regulation by the phosphorylation-dependent factor CREB. Nat Rev *Mol Cell Biol.* **2**:599-609.

McNair K, Davies CH, Cobb SR. (2006) Plasticity-related regulation of the hippocampal proteome. *Eur J Neurosci.* **23**(2):575-80.

McNaughton BL, Douglas RM, Goddard GV. (1978) Synaptic enhancement in fascia dentata: cooperativity among coactive afferents. *Brain Res.* **157**:277-93.

McQuain MK, Seale K, Peek J, Levy S, Haselton FR. (2003) Effects of relative humidity and buffer additives on the contact printing of microarrays by quill pins. *Anal Biochem.* **320**:281-91.

Melzer P, Steiner H. (1997) Stimulus-dependent expression of immediate-early genes in rat somatosensory cortex. *J Comp Neurol.* **380**:145-53.

Merenmies J, Rauvala H. (1990) Molecular cloning of the 18-kDa growth-associated protein of developing brain. *J Biol Chem.* **265**:16721-4.

Meshul CK, Hopkins WF. (1990) Presynaptic ultrastructural correlates of long-term potentiation in the CA1 subfield of the hippocampus. *Brain Res.* **514**:310-9.

Michaelis EK. (1998) Molecular biology of glutamate receptors in the central nervous system and their role in excitotoxicity, oxidative stress and aging. *Prog Neurobiol.* **54**:369-415.
Migaud M, Charlesworth P, Dempster M, Webster LC, Watabe AM, Makhinson M, He Y, Ramsay MF, Morris RG, Morrison JH, O'Dell TJ, Grant SG. (1998) Enhanced long-term potentiation and impaired learning in mice with mutant postsynaptic density-95 protein. *Nature.* **396**:433-9.

Miller MW. (1988) Maturation of rat visual cortex: IV. The generation, migration, morphogenesis, and connectivity of atypically oriented pyramidal neurons. *J Comp Neurol.* **274**:387-405.

Miyashita E, Keller A, and Asanuma H. (1994). Input-output organization of the rat vibrissal motor cortex. *Exp Brain Res* **99**: 223–232

Miyauchi T, Kanekura T, Yamaoka A, Ozawa M, Miyazawa S, Muramatsu T. (1990) Basigin, a new, broadly distributed member of the immunoglobulin superfamily, has strong homology with

both the immunoglobulin V domain and the beta-chain of major histocompatibility complex class II antigen. *J Biochem (Tokyo).* **107**(2):316-23.

Moccia R, Chen D, Lyles V, Kapuya E, E Y, Kalachikov S, Spahn CM, Frank J, Kandel ER, Barad M, Martin KC. (2003) An unbiased cDNA library prepared from isolated Aplysia sensory *Neuron* processes is enriched for cytoskeletal and translational mRNAs. *J Neuroscience.* **23**:9409-17.

Montminy MR, Bilezikjian LM. Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. *Nature.* (1987) Jul 9-15;328(6126):175-8.

Morris RG, Schenk F, Tweedie F, Jarrard LE. (1990) Ibotenate Lesions of Hippocampus and/or Subiculum: Dissociating Components of Allocentric Spatial Learning. *Eur J Neurosci.* **2**:1016-1028.

Mower AF, Liao DS, Nestler EJ, Neve RL, Ramoa AS. (2002) cAMP/Ca2+ response element-binding protein function is essential for ocular dominance plasticity. *J Neurosci.* **22**:2237-45.

Moyner, H., Sprengel, R., Schoepfer, R., Herb, A., Higuchi, M., Lorneli, H., Burnashev, N., Sakmann, B., and Seeburg, P.H. (1992) Heteromeric NMDA receptors: molecular and functional distinction of subtypes. *Science.* **256**: 1217-1221.

Mujumdar RB, Ernst LA, Mujumdar SR, Lewis CJ, Waggoner AS. (1993) Cyanine dye labeling reagents: sulfoindocyanine succinimidyl esters. *Bioconjug Chem.* **4**:105-11.

Mujumdar SR, Mujumdar RB, Grant CM, Waggoner AS. (1996) Cyanine-labeling reagents: sulfobenzindocyanine succinimidyl esters. *Bioconjug Chem.* **7**:356-62.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* **1**:263-73.

Mullis KB, Faloona FA. (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**:335-50.

Murase S, Mosser E, Schuman EM. (2002) Depolarization drives beta-Catenin into *Neuron*al spines promoting changes in synaptic structure and function. *Neuron*. **35**:91-105.

Nairn AC, Shenolikar S. (1992) The role of protein phosphatases in synaptic transmission, plasticity and *Neuron*al development. *Curr Opin Neurobiol*. **2**:296-301.

Naisbitt S, Kim E, Tu JC, Xiao B, Sala C, Valtschanoff J, Weinberg RJ, Worley PF, Sheng M. (1999) Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin. *Neuron*. **23**:569-82.

Naisbitt S, Kim E, Weinberg RJ, Rao A, Yang FC, Craig AM, Sheng M. (1997) Characterization of guanylate kinase-associated protein, a postsynaptic density protein at excitatory synapses that interacts directly with postsynaptic density-95/synapse-associated protein 90. *J Neurosci*. **17**:5687-96.

Nakajima T, Uchida C, Anderson SF, Lee CG, Hurwitz J, Parvin JD, Montminy M. (1997) RNA helicase A mediates association of CBP with RNA polymerase II. *Cell* **90**:1107-12.

Nakanishi, S. (1992) Molecular diversity of glutamate receptors and implications for brain function. *Science*. **258**: 597-603.

Nakic M, Manahan-Vaughan D, Reymann KG, Schachner M. (1998) Long-term potentiation in vivo increases rat hippocampal tenascin-C expression. *J Neurobiol*. **37**:393-404.

Nanda SA, Mack KJ. (2000) Seizures and sensory stimulation result in different patterns of brain derived neurotrophic factor protein expression in the barrel cortex and hippocampus. *Brain Res Mol Brain Res*. **78**:1-14

Naruhashi K, Kadomatsu K, Igakura T, Fan QW, Kuno N, Muramatsu H, Miyauchi T, Hasegawa T, Itoh A, Muramatsu T, Nabeshima T. (1997) Abnormalities of sensory and memory functions in mice lacking Bsg gene. *Biochem Biophys Res Commun*. **236**:733-7.

Purves, Dale; Augustine, George J.; Fitzpatrick, David; Katz, Lawrence C.; LaMantia, Anthony-Samuel; McNamara, James O.; Williams, S. Mark. NeuroScience 2nd ed. (2001) Sunderland: Sinauer Associates, Inc.

Nguyen PV, Kandel ER. (1996) A macromolecular synthesis-dependent late phase of long-term potentiation requiring cAMP in the medial perforant pathway of rat hippocampal slices. *J Neurosci.* **16**:3189-98.

Nguyen T, Sudhof TC. (1997) Binding properties of neuroligin 1 and neurexin 1beta reveal function as heterophilic *Cell* adhesion molecules. *J Biol Chem.* 272:26032-9.

Nicolelis MA, Lin RC, Woodward DJ, Chapin JK. (1993) Dynamic and distributed properties of many-neuron ensembles in the ventral posterior medial thalamus of awake rats. *Proc Natl Acad Sci U S A.* **90**:2212-6.

Nobes CD, Hall A. (1995) Rho, rac and cdc42 GTPases: regulators of actin structures, *Cell* adhesion and motility. *Biochem Soc Trans.* **23**:456-9.

Nowak, L., Bregestovski, P., Ascher, P., Herbert, A., and Prochiantz, A. (1984). Magnesium gates glutamate-activated channels in mouse central neurones. *Nature* **307**:462-465.

Nunez-Garcia J, Mersinias V, Cho KH, Smith CP, Wolkenhauer O. (2003) The statistical distribution of the intensity of pixels within spots of DNA microarrays: what is the appropriate single-value representative? Appl *Bioninform*atics **2**:229-39.

Nusser Z, Lujan R, Laube G, Roberts JD, Molnar E, Somogyi P. (1998) Cell type and pathway dependence of synaptic AMPA receptor number and variability in the hippocampus. *Neuron.* **21**:545-59.

Oleksiak MF, Churchill GA, Crawford DL. (2002) Variation in gene expression within and among natural populations. *Nat Genet.* **32:**261-6.

Pan W. (2002) A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. *Bioninform*atics. **18**:546-54.

Pascale A, Govoni S, Battaini F. (1998) Age-related alteration of PKC, a key enzyme in memory processes: physiological and pathological examples. *Mol Neurobiol.* **16**:49-62.

Patterson SL, Grover LM, Schwartzkroin PA, Bothwell M. (1992) Neurotrophin expression in rat hippocampal slices: a stimulus paradigm inducing LTP in CA1 evokes increases in BDNF and NT-3 mRNAs. *Neuron* **9**:1081-8.

Pavlov I, Voikar V, Kaksonen M, Lauri SE, Hienola A, Taira T, Rauvala H. (2002) Role of heparin-binding growth-associated molecule (HB-GAM) in hippocampal LTP and spatial learning revealed by studies on overexpressing and knockout mice. *Mol Cell Neuroscience.* **20**:330-42.

Pearson G, Robinson F, Beers Gibson T, Xu BE, Karandikar M, Berman K, Cobb MH. (2001) Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions. *Endocr Rev.* **22**:153-83.

Pedrotti B, Islam K. (1996) Dephosphorylated but not phosphorylated microtubule associated protein MAP1B binds to microfilaments. *FEBS Lett.* **388**:131-3.

Peng HB, Ali AA, Dai Z, Daggett DF, Raulo E, Rauvala H. (1995) The role of heparin-binding growth-associated molecule (HB-GAM) in the postsynaptic induction in cultured muscle cells. *J Neuroscience.* **15**:3027-38.

Petersen CC, Grinvald A, Sakmann B. (2003) Spatiotemporal dynamics of sensory responses in layer 2/3 of rat barrel cortex measured in vivo by voltage-sensitive dye imaging combined with whole-Cell voltage recordings and neuron reconstructions. *J Neurosci.* **23**:1298-309.

Petersen CC, Sakmann B. (2000) The excitatory neuronal network of rat layer 4 barrel cortex. *J Neurosci.* **20**:7579-86.

Petersen CC, Sakmann B. (2001) Functionally independent columns of rat somatosensory barrel cortex revealed with voltage-sensitive dye imaging. *J Neurosci.* **21**:8435-46.

Petersen CC. (2003) The barrel cortex--integrating molecular, Cellular and systems physiology. *Pflugers Arch.* **447**:126-34. Epub (2003) Sep 19.

Petersen RS, Diamond ME. (2000) Spatial-temporal distribution of whisker-evoked activity in rat somatosensory cortex and the coding of stimulus location. *J Neurosci.* **20**:6135-43.

Petralia RS, Rubio ME, Wenthold RJ. (1998) Selectivity in the distribution of glutamate receptors in *Neurons*. *Cell Biol Int.* **22**:603-8.

Pham TA, Impey S, Storm DR, Stryker MP. (1999) CRE-mediated gene transcription in neocortical *Neurona*l plasticity during the developmental critical period. *Neuron.* **22**:63-72.

Phillips LL, Reeves TM (2001) Interactive pathology following traumatic brain injury modifies hippocampal plasticity. *Restor Neurol Neuroscience* **19**: 213-235

Picciotto MR, Zoli M, Bertuzzi G, Nairn AC. (1995) Immunochemical localization of calcium/calmodulin-dependent protein kinase I. *Synapse* **20**:75-84.

Porter LL, White EL. (1983) Afferent and efferent pathways of the vibrissal region of primary motor cortex in the mouse. *J Comp Neurol.* **214**:279-89.

Pruijt JF, Fibbe WE, Laterveer L, Pieters RA, Lindley IJ, Paemen L, Masure S, Willemze R, Opdenakker G. (1999) Prevention of interleukin-8-induced mobilization of hematopoietic progenitor cells in rhesus monkeys by inhibitory antibodies against the metalloproteinase gelatinase B (MMP-9). *Proc Natl Acad Sci U S A.* **96**:10863-8.

Pruitt KD, Maglott DR. (2001) RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res.* **29**:137-40.

Purcell AL, Sharma SK, Bagnall MW, Sutton MA, Carew TJ. (2003) Activation of a tyrosine kinase-MAPK cascade enhances the induction of long-term synaptic facilitation and long-term memory in Aplysia. *Neuron* **37**:473-84.

Quackenbush J. (2002) Microarray data normalization and transformation. *Nat Genet* **32**:496-501.

Quinlan EM, Halpain S. (1996) Emergence of activity-dependent, bidirectional control of microtubule-associated protein MAP2 phosphorylation during postnatal development. *J Neurosci.* **16**:7627-37.

Quinn PG. (1993) Distinct activation domains within cAMP response element-binding protein (CREB) mediate basal and cAMP-stimulated transcription. *J Biol Chem.* **268**:16999-7009.

Rajeevan MS, Ranamukhaarachchi DG, Vernon SD, Unger ER. (2001) Use of real-time quantitative PCR to validate the results of cDNA array and differential display PCR technologies. *Methods*. **25**:443-51.

Ramon Y Cajal S. (1952) Structure and connections of neurons. *Bull Los Angel Neuro Soc.* **17**:5-46.

Rauvala H, Pihlaskari R. (1987) Isolation and some characteristics of an adhesive factor of brain that enhances neurite outgrowth in central *Neurons*. *J Biol Chem.* **262**:16625-35.

Rauvala H. (1989) An 18-kd heparin-binding protein of developing brain that is distinct from fibroblast growth factors. *EMBO J.* **8**:2933-41.

Rebhan, M., Chalifa-Caspi, V., Prilusky, J., Lancet, D. (1998) GeneCards: a novel functional genomics compendium with automated data mining and query reformulation support. *Bioinformatics*. **14**:656–664

Redmond L, Ghosh A. (2001) The role of Notch and Rho GTPase signaling in the control of dendritic development. *Curr Opin Neurobiol.* **11**:111-7.

Rehfuss RP, Walton KM, Loriaux MM, Goodman RH. (1991) The cAMP-regulated enhancer-binding protein ATF-1 activates transcription in response to cAMP-dependent protein kinase A. *J Biol Chem.* **266**:18431-4.

Reid SN, Daw NW. (1995) Dark-rearing changes dendritic microtubule-associated protein 2 (MAP2) but not subplate *Neurons* in cat visual cortex. *J Comp Neurol.* **359**:38-47.

Rema V, Armstrong-James M, Ebner FF. (1998) Experience-dependent plasticity of adult rat S1 cortex requires local NMDA receptor activation. *J Neurosci.* **18**:10196-206.

Renehan WE, Munger BL. (1986) Degeneration and regeneration of peripheral nerve in the rat trigeminal system. I. Identification and characterization of the multiple afferent innervation of the mystacial vibrissae. *J Comp Neurol.* **246**:129-45.

Reyes A, Sakmann B. (1999) Developmental switch in the short-term modification of unitary EPSPs evoked in layer 2/3 and layer 5 pyramidal neurons of rat neocortex. *J Neurosci.* **19**:3827-35.

Rice FL, Gomez C, Barstow C, Burnet A, Sands P. (1985) A comparative analysis of the development of the primary somatosensory cortex: interspecies similarities during barrel and laminar development. *J Comp Neurol.* **236**:477-95.

Rice FL, van der Loos H (1977) Development of the barrels and barrel field in the somatosensory cortex of the mouse. *J Comp Neurol* **171**:545–560.

Richardson CL, Tate WP, Mason SE, Lawlor PA, Dragunow M, Abraham WC. (1992) Correlation between the induction of an immediate early gene, zif/268, and long-term potentiation in the dentate gyrus. *Brain Res.* **580**:147-54.

Ridley AJ, Paterson HF, Johnston CL, Diekmann D, Hall A. (1992) The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell.* **70**:401-10.

Roberts CJ, Nelson B, Marton MJ, Stoughton R, Meyer MR, Bennett HA, He YDD, Dai HY, Walker WL, Hughes TR, Tyers M, Boone C, Friend SH (2000) Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. *Science* **287**: 873-880.

Robertson HA. (1992) Immediate-early genes, *Neurona*l plasticity, and memory. *Biochem Cell Biol.* **70**:729-37.

Rocke DM, Durbin B. (2001); A model for measurement error for gene expression arrays. *J Comput Biol.* **8**:557-69.

Rosahl TW, Geppert M, Spillane D, Herz J, Hammer RE, Malenka RC, Sudhof TC. (1993) Short-term synaptic plasticity is altered in mice lacking synapsin I. *Cell.* **75**:661-70.

Rosen LB, Ginty DD, Weber MJ, Greenberg ME. (1994) Membrane depolarization and calcium influx stimulate MEK and MAP kinase via activation of Ras. *Neuron.* **12**:1207-21.

Ryan TA, Joiner BL (1976). Normal Probability Plots and Tests for Normality. Technical Report, Statistics Department, The Pennsylvania State University. (Available from Minitab Inc.)

Sambrook J, Fritsch E F, Maniatis T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanchez C, Diaz-Nido J, Avila J. (2000) Phosphorylation of microtubule-associated protein 2 (MAP2) and its relevance for the regulation of the *Neuron*al cytoskeleton function. *Prog Neurobiol.* **61**:133-68.

Sattilaro RF. (1986) Interaction of microtubule-associated protein 2 with actin filaments. *Biochemistry.*;**25**(8):(2003)-9.

Schena M, Shalon D, Davis RW, Brown PO. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science.* **270**:467-70.

Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW. (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci U S A.* **93**:10614-9.

Schermer, M.J. (1999) Confocal scanning microscopy in microarray detection. In: Schena, M. (Ed.) DNA Microarrays: A Practical Approach *Oxford University Press* pp.17-42.

Schlaggar BL, Fox K, and O'Leary DD. (1993) Postsynaptic control of plasticity in developing somatosensory cortex. *Nature* **364**: 623-626,

Scoville WB, Milner B. (1957) Loss of recent memory after bilateral hippocampal lesions. *J Neurochem.* **20**:11-21.

Senft SL, Woolsey TA. (1991) Growth of thalamic afferents into mouse barrel cortex. *Cereb Cortex.* **1**:308-35.

Shapiro, S. S., Wilk, M. B. (1965). Biometrika. **52**:591-611.

Sheng M. (2001) Molecular organization of the postsynaptic specialization. *Proc Natl Acad Sci U S A.* **98**:7058-61.

Shi SH, Hayashi Y, Petralia RS, Zaman SH, Wenthold RJ, Svoboda K, Malinow R. (1999) Rapid spine delivery and redistribution of AMPA receptors after synaptic NMDA receptor activation. *Science.* **284**:1811-6.

Shimamura M, Garcia JM, Prough DS, Dewitt DS, Uchida T, Shah SA, Avila MA, Hellmich HL. (2005) Analysis of long-term gene expression in neurons of the hippocampal subfields following traumatic brain injury in rats. *Neuroscience*. **131**(1):87-97.

Sidak, Z. (1967) Rectangular confidence regions for the means of multivariate normal distribution. *J. Am. Statist. Assoc.* **62**:626–633

Si K, Giustetto M, Etkin A, Hsu R, Janisiewicz AM, Miniaci MC, Kim JH, Zhu H, Kandel ER. (2003) A *Neuron*al isoform of CPEB regulates local protein synthesis and stabilizes synapse-specific long-term facilitation in aplysia. *Cell*. **115**:893-904.

Silva AJ, Wang Y, Paylor R, Wehner JM, Stevens CF, Tonegawa S. (1992) Alpha calcium/calmodulin kinase II mutant mice: deficient long-term potentiation and impaired spatial learning. *Cold Spring Harb Symp Quant Biol.* **57**:527-39.

Sim AT. (1991) The regulation and function of protein phosphatases in the brain. *Mol Neurobiol.* **5**:229-46.

Simon R, Radmacher MD, Dobbin K. (2002) Design of studies using DNA microarrays. *Genet Epidemiol.* **23**:21-36.

Simons DJ and Land PW. (1987) Early experience of tactile stimulation influences organization of somatic sensory cortex. *Nature* **326**: 694–697,

Simons DJ, Carvell GE. (1989) Thalamocortical response transformation in the rat vibrissa/barrel system. *J Neurophysiol.* **61**:311-30.

Simons DJ, Woolsey TA. (1984) Morphology of Golgi-Cox-impregnated barrel neurons in rat SmI cortex. *J Comp Neurol.* **230**:119-32.

Simons DJ. (1978) Response properties of vibrissa units in rat SI somatosensory neocortex. *J Neurophysiol.* **41**:798-820.

Slonim DK. (2002) From patterns to pathways: gene expression data analysis comes of age. *Nat Genet.* **32**:502-8.

Soille, P. (1999) *Morphological Image Analysis.* Berlin: Springer-Verlag.

Sollner T, Bennett MK, Whiteheart SW, Scheller RH, Rothman JE. (1993) A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell.* **75**:409-18.

Southern EM. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol.* **98**:503-17.

Speed T. (2003) Statistical Analysis of Gene Expression Microarray Data. *Chapman and Hall / CRC*

Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, Futcher B. (1998) Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol Biol Cell.* **9**:3273-97.

Sprengel R, Suchanek B, Amico C, Brusa R, Burnashev N, Rozov A, Hvalby O, Jensen V, Paulsen O, Andersen P, Kim JJ, Thompson RF, Sun W, Webster LC, Grant SG, Eilers J, Konnerth A, Li J, McNamara JO, Seeburg PH. (1998) Importance of the intra*Cell*ular domain of NR2 subunits for NMDA receptor function in vivo. *Cell.* **92**:279-89.

Squire LR, Cohen NJ, Zouzounis JA. (1984); Preserved memory in retrograde amnesia: sparing of a recently acquired skill. *Neuropsychologia.* **22:**145-52.

Stanton, P. K. & Sarvey, J. M. (1984). *J. Neurosci.* **4**: 3080-3088

Steigerwald F, Schulz TW, Schenker LT, Kennedy MB, Seeburg PH, Kohr G. (2000) C-Terminal truncation of NR2A subunits impairs synaptic but not extrasynaptic localization of NMDA receptors. *J Neurosci.* **20**:4573-81.

Sudhof TC. (2004) The synaptic vesicle cycle. *Annu Rev Neurosci.* **27**:509-47

Sugita S, Saito F, Tang J, Satz J, Campbell K, Sudhof TC. (2001) A stoichiometric complex of neurexins and dystroglycan in brain. *J Cell Biol.* **154**:435-45.

Sun P, Lou L, Maurer RA. (1996) Regulation of activating transcription factor-1 and the cAMP response element-binding protein by Ca2+/calmodulin-dependent protein kinases type I, II, and IV. *J Biol Chem.* **271**:3066-73.

Suzuki T, Higgins PJ, Crawford DR. (2000) Control selection for RNA quantitation. *Biotechniques*. **29**:332-7.

Svitkina TM, Verkhovsky AB, McQuade KM, Borisy GG. (1997) Analysis of the actin-myosin II system in fish epidermal keratocytes: mechanism of *Cell* body translocation. *J Cell Biol.* **139**:397-415.

Sweatt JD (2004) Mitogen-activated protein kinases in synaptic plasticity and memory. *Curr Opin Neurobiol.* **3**:311-7.

Szklarczyk A, Lapinska J, Rylski M, McKay RD, Kaczmarek L. (2002) Matrix metalloproteinase-9 undergoes expression and activation during dendritic remodeling in adult hippocampus. *J Neuroscience.* **22**:920-30.

Takeda A, Onodera H, Sugimoto A, Itoyama Y, Kogure K, Rauvala H, Shibahara S. (1995) Induction of heparin-binding growth-associated molecule expression in reactive astrocytes following hippocampal *Neuron*al injury. *NeuroScience.* **68**:57-64.

Takei Y, Teng J, Harada A, Hirokawa N. (2000) Defects in axonal elongation and *Neuron*al migration in mice with disrupted tau and map1b genes. *J Cell Biol.* **150**:989-1000.

Takumi Y, Matsubara A, Rinvik E, Ottersen OP. (1999) The arrangement of glutamate receptors in excitatory synapses. *Ann N Y Acad Sci.* ;**868**:474-82.

Tanaka TS, Jaradat SA, Lim MK, Kargul GJ, Wang X, Grahovac MJ, Pantano S, Sano Y, Piao Y, Nagaraja R, Doi H, Wood WH 3rd, Becker KG, Ko MS. (2000) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc Natl Acad Sci U S A.* **97**:9127-32.

Tang L, Hung CP, Schuman EM.(1998) A role for the cadherin family of cell adhesion molecules in hippocampal long-term potentiation. *Neuron.* **6**:1165-75.

Tao H, Bausch C, Richmond C, Blattner FR, Conway T. (1999) Functional genomics: expression analysis of Escherichia coli growing on minimal and rich media. *J Bacteriol.* **181**:6425-40.

Taylor SS, Buechler JA, Yonemoto W (1990) cAMP-dependent protein kinase: framework for a diverse family of regulatory enzymes. *Annu Rev Biochem* **59**:971-1005

Teng J, Takei Y, Harada A, Nakata T, Chen J, Hirokawa N. (2001) Synergistic effects of MAP2 and MAP1B knockout in *Neuron*al migration, dendritic outgrowth, and microtubule organization. *J Cell Biol.* **155**:65-76.

ter Linde JJ, Liang H, Davis RW, Steensma HY, van Dijken JP, Pronk JT. (1999) Genome-wide transcriptional analysis of aerobic and anaerobic chemostat cultures of Saccharomyces cerevisiae. *J Bacteriol.* **181**(24):7409-13.

Thomson S, Clayton AL, Hazzalin CA, Rose S, Barratt MJ, Mahadevan LC. (1999) The nucleosomal response associated with immediate-early gene induction is mediated via alternative MAP kinase cascades: MSK1 as a potential histone H3/HMG-14 kinase. *EMBO J.* **18**:4779-93.

Tiedge H, Brosius J. (1996) Translational machinery in dendrites of hippocampal *Neuron*s in culture. *J Neuroscience.* 16:7171-81.

Togashi H, Abe K, Mizoguchi A, Takaoka K, Chisaka O, Takeichi M. (2002) Cadherin regulates dendritic spine morphogenesis. *Neuron.* 35:77-89.

Togel M, Wiche G, Propst F. (1998) Novel features of the light chain of microtubule-associated protein MAP1B: microtubule stabilization, self interaction, actin filament binding, and regulation by the heavy chain. *J Cell Biol.* **143**:695-707.

Tokuda M, Hatase O. (1998) Regulation of *Neuron*al plasticity in the central nervous system by phosphorylation and dephosphorylation. *Mol Neurobiol.* **17**:137-56

Toni N, Buchs PA, Nikonenko I, Bron CR, Muller D. (1999) LTP promotes formation of multiple spine synapses between a single axon terminal and a dendrite. *Nature* **402**:421-5.

Tran PH, Peiffer DA, Shin Y, Meek LM, Brody JP, Cho KW. (2002) Microarray optimizations: increasing spot accuracy and automated identification of true microarray signals. *Nucleic Acids Res* **30**:e54.

Tseng GC, Oh MK, Rohlin L, Liao JC, Wong WH. (2001) Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res.* **29**:2549-57.

Tsuyama S, Terayama Y, Matsuyama S. (1987) Numerous phosphates of microtubule-associated protein 2 in living rat brain. *J Biol Chem.* **262**:10886-92

Tu JC, Xiao B, Naisbitt S, Yuan JP, Petralia RS, Brakeman P, Doan A, Aakalu VK, Lanahan AA, Sheng M, Worley PF. (1999) Coupling of mGluR/Homer and PSD-95 complexes by the Shank family of postsynaptic density proteins. *Neuron.* **23**:583-92.

Tusher VG, Tibshirani R, Chu G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A.* **98**:5116-21.

Van Aelst L, D'Souza-Schorey C. (1997) Rho GTPases and signaling networks. *Genes Dev.* **11**:2295-322.

Van der Loos H, Woolsey TA. (1973) Somatosensory cortex: structural alterations following early injury to sense organs. *Science.* **179**:395-8.

Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH. (1990) Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci U S A.* **87**:1663-7.

Veinante P, Deschenes M. (1999) Single- and multi-whisker channels in the ascending projections from the principal trigeminal nucleus in the rat. *J Neurosci.* **19**:5085-95.

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. (1995) Serial analysis of gene expression. *Science.* **270**:484-7.

Verdoorn T.A., Burnashev N., Monyer H., Seeburg P.H. & Sakmann B. (1991) Structural determinants of ion flow through recombinant glutamate receptor channels. *Science* **252**:1715-1718.

Vincent SB. (1912) The function of the vibrissae in the behavior of the white rat. *Behav Monogr* **1**:7–81

Voets T, Toonen RF, Brian EC, de Wit H, Moser T, Rettig J, Sudhof TC, Neher E, Verhage M. (2001) Munc18-1 promotes large dense-core vesicle docking. *Neuron.* **31**:581-91.

Vossler MR, Yao H, York RD, Pan MG, Rim CS, Stork PJ. (1997) cAMP activates MAP kinase and Elk-1 through a B-Raf- and Rap1-dependent pathway. *Cell.* **89**:73-82.

Vyavahare N, Jones PL, Tallapragada S, Levy RJ. (2000) Inhibition of matrix metalloproteinase activity attenuates tenascin-C production and calcification of implanted purified elastin in rats. *Am J Pathol.* **157**:885-93.

Waite PM. (1972) Unitary responses in the rat's ventro-basal complex to ramp-shaped movements of the whiskers. *J Physiol.* **222**:59P-60P.

Walaas SI, Nairn AC. (1989); Multisite phosphorylation of microtubule-associated protein 2 (MAP-2) in rat brain: peptide mapping distinguishes between cyclic AMP-, calcium/calmodulin-, and calcium/phospholipid-regulated phosphorylation mechanisms. *J Mol Neurosci.* 1:117-27.

Wallace H, Fox K. (1999) The effect of vibrissa deprivation pattern on the form of plasticity induced in rat barrel cortex. *Somatosens Mot Res.* **16**:122-38.

Wallace H, Glazewski S, Liming K, Fox K. (2001) The role of cortical activity in experience-dependent potentiation and depression of sensory responses in rat barrel cortex. *J Neuroscience* **21**:3881-94.

Wanaka A, Carroll SL, Milbrandt J. (1993) Developmentally regulated expression of pleiotrophin, a novel heparin binding growth factor, in the nervous system of the rat. *Brain Res.* **72**:133-44.

Wayman GA, Tokumitsu H, Soderling TR. (1997) Inhibitory cross-talk by cAMP kinase on the calmodulin-dependent protein kinase cascade. *J Biol Chem.* **272**:16073-6.

Wayman GA, Wei J, Wong S, Storm DR. (1996) Regulation of type I adenylyl cyclase by calmodulin kinase IV in vivo. *Mol Cell Biol.* **16**:6075-82.

Welker C. (1976) Receptive fields of barrels in the somatosensory neocortex of the rat. *J Comp Neurol.* **166**:173-89.

Welker WI. (1964) .Analysis of sniffing of the albino rat. *Behaviour* **12**: 223–244,

Weller WL. (1972) Barrels in somatic sensory neocortex of the marsupial Trichosurus vulpecula (brush-tailed possum). *Brain Res.* **43**:11-24.

Weller WL, Johnson JI. (1975) Barrels in cerebral cortex altered by receptor disruption in newborn, but not in five-day-old mice (Cricetidoe and Muridae). *Brain Res.* **83**:504-8.

Wenzel J, Desmond NL, Levy WB. (1993) Somatic ribosomal changes induced by long-term potentiation of the perforant path-hippocampal CA1 synapses. *Brain Res.* **619**:331-3.

Wenzel J, Kammerer E, Kirsche W, Matthies H, Wenzel M. (1980) Electron microscopic and morphometric studies on synaptic plasticity in the hippocampus of the rat following conditioning. *J Hirnforsch.* **21**:647-54.

West AE, Chen WG, Dalva MB, Dolmetsch RE, Kornhauser JM, Shaywitz AJ, Takasu MA, Tao X, Greenberg ME. (2001) Calcium regulation of neuronal gene expression. *Proc Natl Acad Sci U S A.* **98**:11024-31.

Westfall P.H., and Young S.S. (1993) Resampling-Based Multiple Testing. Examples and Methods for p-Values Adjustment. *John Wiley & Sons, New York.*

Wheeler DL, Chappey C, Lash AE, Leipe DD, Madden TL, Schuler GD, Tatusova TA, Rapp BA. (2000) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **28**:10-4

Wiche G, Oberkanins C, *Himmler A. (1991); Molecular structure and function of microtubule-associated proteins. Int Rev Cytol.* **124**:217-73.

Wigstrom H, Gustafsson B. (1986); Postsynaptic control of hippocampal long-term potentiation. *J Physiol (Paris).* **81**:228-36.

Wille H, Mandelkow EM, Dingus J, Vallee RB, Binder LI, Mandelkow E. (1992) Domain structure and antiparallel dimers of microtubule-associated protein 2 (MAP2). *J Struct Biol.* **108**:49-61.

Wille H, Mandelkow EM, Mandelkow E. (1992) The juvenile microtubule-associated protein MAP2c is a rod-like molecule that forms antiparallel dimers. *J Biol Chem.* **267**:10737-42.

Wineski LE. (1985) Facial morphology and vibrissal movement in the golden hamster. *J Morphol.* **183**:199-217.

Wineski, LE. (1983) Movements of the cranial vibrissae in the golden hamster (*Mesocritus auratus*). *J. Zool. (Lond.)* **200**: 261-280,

Wisden W, Errington ML, Williams S, Dunnett SB, Waters C, Hitchcock D, Evan G, Bliss TV, Hunt SP. (1990) Differential expression of immediate early genes in the hippocampus and spinal cord. *Neuron.* **4**:603-14.

Woo Y, Affourtit J, Daigle S, Viale A, Johnson K, Naggert J, Churchill G. (2004) A comparison of cDNA, oligonucleotide, and Affymetrix GeneChip gene expression microarray platforms. *J Biomol Tech.* **15**:276-84.

Woolsey TA, Van der Loos H. (1970) The structural organization of layer IV in the somatosensory region (SI) of mouse cerebral cortex. The description of a cortical field composed of discrete cytoarchitectonic units. *Brain Res.* **17**:205-42.

Woolsey TA, Wann JR. (1976) Areal changes in mouse cortical barrels following vibrissal damage at different postnatal ages. *J Comp Neurol.* **170**:53-66.

Woolsey TA, Welker C, Schwartz RH. (1975) Comparative anatomical studies of the SmL face cortex with special reference to the occurrence of "barrels" in layer IV. *J Comp Neurol.* **164**:79-94.

Wu GY, Deisseroth K, Tsien RW. (2001) Activity-dependent CREB phosphorylation: convergence of a fast, sensitive calmodulin kinase pathway and a slow, less sensitive mitogen-activated protein kinase pathway. *Proc Natl Acad Sci U S A.* **98**:2808-13.

Xia J, Zhang X, Staudinger J, Huganir RL. (1999) Clustering of AMPA receptors by the synaptic PDZ domain-containing protein PICK1. *Neuron.* **22**:179-87.

Xia Z, Storm DR. (1997) Calmodulin-regulated adenylyl cyclases and neuromodulation. *Curr Opin Neurobiol.* **7**:391-6.

Xiao Y, Segal MR, Rabert D, Ahn AH, Anand P, Sangameswaran L, Hu D, Hunt CA. (2002) Assessment of differential gene expression in human peripheral nerve injury. *BMC Genomics.* **3**:28.

Xing L, Quinn PG. (1994) Three distinct regions within the constitutive activation domain of cAMP regulatory element-binding protein (CREB) are required for transcription activation. *J Biol Chem.* **269**:28732-6.

Yamada H, Saito F, Fukuta-Ohi H, Zhong D, Hase A, Arai K, Okuyama A, Maekawa R, Shimizu T, Matsumura K. (2001) Processing of beta-dystroglycan by matrix metalloproteinase disrupts the link between the extracellular matrix and cell membrane via the dystroglycan complex. *Hum Mol Genet.* **10**:1563-9.

Yang IV, Chen E, Hasseman JP, Liang W, Frank BC, Wang S, Sharov V, Saeed AI, White J, Li J, Lee NH, Yeatman TJ, Quackenbush J. (2002) Within the fold: assessing differential expression measures and reproducibility in microarray assays. *Genome Biol.* **3**:research0062.

Yang YH (a), Buckley MJ, Speed TP. (2001) Analysis of cDNA microarray images. *Brief Bioinform.* **2**:341-9.

Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30**:e15.

Yang YH, Speed T. (2002) Design issues for cDNA microarray experiments. *Nat Rev Genet.* **3**:579-88.

Yang, YH. (b); Dudoit, SD.; Luu, P.; Speed, TP. (2001) Normalization for cDNA Microarray Data. In SPIE BioE. .

Yauk CL, Berndt ML, Williams A, Douglas GR. (2004) Comprehensive comparison of six microarray technologies. *Nucleic Acids Res.* **32**:e124.

Yin JC, Wallach JS, Del Vecchio M, Wilder EL, Zhou H, Quinn WG, Tully T. (1994) Induction of a dominant negative CREB transgene specifically blocks long-term memory in Drosophila. *Cell.* **79**:49-58.

Yu H, Chao J, Patek D, Mujumdar R, Mujumdar S, Waggoner AS. (1994) Cyanine dye dUTP analogs for enzymatic labeling of DNA probes. *Nucleic Acids Res.* **22**:3226-32

Yue H, Eastman PS, Wang BB, Minor J, Doctolero MH, Nuttall RL, Stack R, Becker JW, Montgomery JR, Vainer M, Johnston R. (2001) An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression. *Nucleic Acids Res.* **29**:E41-1.

Zaccaria ML, Perrone-Capano C, Melucci-Vigo G, Gaeta L, Petrucci TC, Paggi P. (2001) Differential regulation of transcripts for dystrophin Isoforms, dystroglycan, and alpha3AChR subunit in mouse sympathetic ganglia following postganglionic nerve crush. *Neurobiol Dis.* **8**:513-24.

Zhang JW, Deb S, Gottschall PE. (1998) Regional and differential expression of gelatinases in rat brain after systemic kainic acid or bicuculline administration. *Eur J Neuroscience.* **10**:3358-68.

Zhu JJ, Connors BW. (1999) Intrinsic firing patterns and whisker-evoked synaptic responses of neurons in the rat barrel cortex. *J Neurophysiol.* **81**:1171-83.

Zien, A., Fluck, J., Zimmer, R., Lengauer, T. (2003) Microarrays: how many do you need? *J. Comput. Biol.*, **10**:653–667