

**Robotic platforms for large-scale analysis of gene expression on
tissue sections by means of *in situ* hybridization**

Von dem Fachbereich Chemie der Universität Hannover
zur Erlangung des Grades eines
Doctors der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation
von
M.Sc. Biol. Murat Burak Yaylaoglu
geboren am 07.Juli.1971 in Ankara
- 2003 -

This study was carried out under the supervision of Prof. Dr. G. Eichele at the Max Planck Institute of Experimental Endocrinology in Hannover, between February 2001 and May 2003.

Referent : Prof. Dr. W. H. Mueller
Medizinische Hochschule
Hannover

Korreferent : Prof. Dr. G. Eichele
Max Planck Institute of Experimental Endocrinology,
Hannover

Tag der Promotion : 15 July.2002

Acknowledgements

I will most probably miss out on properly expressing my appreciation or leaving out a lot of people I should thank, so I want to start by apologizing to these people for I presume keeping this longer than one page is not feasible.

Thanking Prof. Eichele will be very complicated and futile, for words will not be enough to express my gratitude. To have worked with one of the best scientists on the globe puts a lot of pressure on you, especially if you try to learn and gain from the level of perfectionism that has been acquired through years. Everything has run very smoothly with your endless encouragement and guidance, thank you.

I would very much like to thank Prof. Müller for his kindness and wisdom. I think without the confidence he has made me feel I would not have been able to deal with my personal insecurity or with Turkish formalities (obligations of military service). Thank you for your guidance.

I would very much like to thank Kornelia Maslo and Klaus Ebert for sharing their experience with the ISH technology and making my first days in the institute much better. Reiner Psala for always being there. I would like to thank the workshop; Sigmar Falkenhagen and especially Uwe Herzig who has modified the Tecan platform, I am really wondering what else Uwe will come up with. I would like to thank Barbara Fischer for the excellent technical assistance. I would like to thank the administration, especially Carsten Gottschalk for patiently helping with formalities. I would likewise like to thank Valerie Ashe. I would like to thank our collaborators in Switzerland and Italy.

Melina Schuh and Andrew Titmus were wonderful people to supervise and work with, its amazing how much you can learn while trying to explain something. I would like to thank Tarvo Thamm, Lars Geffers, Qihong Jiang, Roland Rabeler and Theo Papadopoulos (Greece) for their friendship. I would like to thank my colleagues in the institute Dr. Henrik Oster, Carsten Möller, Axel Visel, Dr. Gonzalo Alvarez-Bolado, Dr. Xunlei Zhou, Dr. Christina Cadenas, Dr. Miki Tsukada, Markus Hükel, Marei Warnecke and Judit Oldekamp for their questions, fruitful discussion and scientific support. I must thank Christina Zwingman Nadine Najokat and Heike Krause for their help when I ran out of tissue sections and mice. Finally I would like to thank every member of the Max Planck Institute in Hannover.

Mum, dad heres my doctor thesis. To my family (Kerim, Umut, Ismail, Ralph, Safak, Akillilar, Yenidunyalari).

Zusammenfassung

Die Genome von Mensch und Maus sind jetzt sequenziert. Diese bemerkenswerte Errungenschaft drängt uns nun zum nächsten Schritt: die Entschlüsselung der Genfunktionen. Um die Funktion eines Gens zu verstehen, ist es notwendig, die Genexpression und in den Zellen zu studieren. Wir haben ein einzigartiges und leistungsfähiges Werkzeug entwickelt, genannt GenePaint, mit dem wir die Genexpression in Gewebeschnitten erfassen und dokumentieren können. Im Einzelnen haben wir entwickelt und gebaut:

- einen Roboter für die automatisierte *In situ* Hybridisierung von Gewebeschnitten im Hochdurchsatzverfahren, der Daten mit zellulärer Auflösung erbringt,
- ein automatisiertes Hochdurchsatz-Dia-Abtastungsmikroskop zur Digitalisierung der Genexpressionsmuster,
- sowie eine Datenbank (GenePaint.org), die die kommentierten Genexpressionsmuster zugänglich macht.

In dieser Studie wurden einzelne Bestandteile der GenePaint-Technologie geprüft, sowohl Hardware als auch Reagenzien, und mit herkömmlichen Niedrigdurchsatz-Verfahren verglichen. Bei herkömmlichen Methoden zur Expressionsanalyse, wie Northern blot, sind die Informationen über die räumliche Verteilung der Genabschrift (des Transkripts) minimal. Darüber hinaus kann sogar eine starke aber räumlich begrenzte Expression eines Gens ein schwaches oder nicht erkennbares Signal im Northern blot ergeben, wenn die Expression auf eine kleine Zellpopulation innerhalb des Gewebes beschränkt ist. Mit unserem ISH-Verfahren werden Expressionsmuster in einer einzelnen Zelle sichtbar gemacht; so steht ein Nachweis hoher Empfindlichkeit für die Hochdurchsatz-Analyse der Genexpression zur Verfügung.

Um die Leistungsfähigkeit dieses integrierten Systems zu veranschaulichen, wurden alle bekannten Mausorthologe der menschlichen Gene des Chromosoms 21 (HC21) analysiert. Die Resultate der Expressionsanalyse sind jetzt öffentlich auf der Website www.tigem.it/ch21exp/ zugänglich und seit kurzem auch in der GenePaint-Datenbank unter www.genepaint.org vorhanden. Diese Studie zeigt, daß es jetzt möglich ist, die Expressionsmuster aller Gene des Mausgenoms zu bestimmen. Folglich kann diese Technologie für Genom-umfassende Studien der Genexpression verwendet werden, und

sie fügt eine neue Facette zur Ermittlung der Genfunktion hinzu. Wir werden nicht nur Einblick in verschiedene pathologische Störungen gewinnen, sondern diese Technik ermöglicht es uns außerdem, die Genregulation und die Netzwerke der Gene auf der Ebene des Genoms zu verstehen.

Schlagworte: Menschliches Chromosom 21, Expressionsmuster-Analyse, Hochdurchsatz-*in situ*-Hybridisierung.

Abstract

The human and mouse genomes are now sequenced. As a result of this remarkable progress, the need to decipher gene function has become even more pressing. To advance understanding of gene function a necessary step is to study gene expression at tissue and cellular levels. We have developed a unique and powerful tool termed GenePaint to study and document gene expression on tissue sections. Specifically we have:

- Constructed instrumentation for high-throughput, automated *in situ* hybridization on tissue sections that yield data with cellular resolution.
- Assembled an automated high-throughput slide scanning microscope that digitizes tissue sections and their gene expression patterns.
- Established a database (GenePaint.org) that makes accessible annotated gene expression patterns.

In this study individual components of GenePaint hardware and chemistry were thoroughly tested and were compared with conventional low-throughput procedures. When using conventional methods such as northern hybridization to analyse expression, the information on the spatial distribution of transcript is minimal. In addition, even strong but localized expression of a gene may result in a weak or undetectable signal in northern blots if expression is restricted to a small population of cells within the tissue. Using our ISH and image acquisition system, expression patterns are revealed at a single cell resolution; thus a high sensitivity assay is provided for high-throughput analysis of gene expression.

To illustrate the power of this integrated system all known mouse orthologues of the human chromosome 21 (HC21) genes were analyzed. The results of the expression analysis are now publicly available at the website www.tigem.it/ch21exp/ and having recently been uploaded onto the GenePaint database, are also available at www.genepaint.org. This study demonstrates that it is now possible to determine the expression pattern of all genes of the mouse genome. Therefore this technology may be used for genome-wide studies of gene expression, adding a new facet to the determination of function. In turn this will not only allow us to gain insight into various disorders, but will enable us to understand gene regulation and gene networks at a genomic level.

Keywords: Human chromosome 21, expression pattern analysis, high-throughput *in situ* hybridization.

Contents

Contents	8
Foreword.....	10
Abbreviations.....	11
I. Introduction	12
1.1 Large scale expression analysis.....	12
1.2 Diseases related to chromosome 21	15
1.3 Relationships existing between expression patterns and the organization of the genome....	18
II. Materials and Methods	20
2.1 Materials.....	20
2.2 Methods.....	22
2.2.1 Isolation of nucleic acids	22
2.2.2 Determination of the nucleic acid concentration	23
2.2.3 Restriction of DNA.....	23
2.2.4 Ligation of DNA fragments.....	23
2.2.5 Transformation of competent bacterial cells	24
2.2.6 Preparation of LB medium and LB agar plates	24
2.2.7 Agarose gel electrophoresis.....	25
2.2.8 Polymerase Chain Reaction (PCR).....	25
2.2.9 Non-radioactive dye terminator cycle sequencing.....	25
2.2.10 Reverse transcription	26
2.2.11 Quantitative PCR and the determination of detection sensitivity	27
2.2.12 cDNA's representing the mouse orthologues of the HC21 genes	30
2.2.13 RNA detection by automated <i>in situ</i> hybridization: instrumentation for high throughput gene expression analysis	37
III. Results.....	51
3.1 Hapten labeled non-radioactive RNA <i>in situ</i> hybridization	51
3.1.1. Effect of Tyramide Signal Amplification and low background.....	51
3.1.2 Comparison of radioactive and non-radioactive ISH.....	53
3.1.3 Quantification of the copy number of Dscr3 transcripts in extracted mRNA from P7 brain sample using real-time PCR.....	55
3.1.4 Improved ISH performance using optimized probes	58
3.2 Gene expression atlas of the mouse orthologues of the majority of human chromosome 21 genes	59
3.2.1 Statistics	62
3.2.2 Genes expressed within the brain	64
3.2.3 Genes expressed within the heart	71

3.2.4 Genes expressed within the limbs	72
3.2.5 Genes expressed within the digestive tract	74
3.2.6 Genes expressed within the thymus.....	75
3.2.7 Genes expressed within the pancreas	76
3.2.8 Phenotypic comparison of Down with Williams Syndrome.....	76
3.2.9 Implications of clustering for regulation of expression	78
3.3 GenePaint database	83
IV. Discussion.....	85
4.1 Evaluation of the non-radioactive <i>in situ</i> hybridization protocol	85
4.1.1 Quality of in <i>in situ</i> hybridization data depends on the source of the DNA template ..	88
4.2 Analysis of the HC21 orthologue mouse gene expression data	89
4.2.1 Down syndrome	90
4.2.2 Phenotypic comparison of Down with Williams Syndrome.....	93
4.2.3 Implications for clustering of expression patterns, for shared regulation.....	93
4.3 Future directions.....	95
V. Conclusion	97
VI. References	98
VII. Appendix	103
7.1 Brief introduction to www.tigem.it/ch21exp/	103
7.2 Brief introduction to www.genepaint.org	107
Curriculum vitae	111

Foreword

To have worked on such a thesis project, has been a remarkable experience for a lot of reasons but the greatest challenge was to actually absorb and interpret the massive amount of data that was being produced. Collection of a lot of experimental data, which was demanding, is usually the case when working with robotics.

This thesis thus presents the potential of the power of automated *in situ* hybridization, how the data are produced, organized and stored for optimal retrieval. It is also illustrated how the data can be accessed on web-based databases. Last but not least I discuss some of the more obvious scientific implications of my work and how this may shape some of the current views held in the field of genomics.

The need for an approach used in the present work arises from the vast amount of data that needs to be examined as biology moves into the post-genomic era.

Abbreviations

CPM	Counts per minutes
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide
DS	Down syndrome
DIG	Digoxigenin
DTT	Dithiothreitol
E14.5	Embryonic day 14.5
EDTA	Ethylene diamine tetraacetic acid
HC21	Human chromosome 21
ISH	<i>In situ</i> hybridization
IPTG	Isopropyl- β -thiogalactopyranoside
LB	Luria-Bertani medium
LCRs	Locus control regions
mL	Milliliter
min	Minutes
NaAc	Sodium acetate
NBT	Nitro-blue tetrazolium
NTE	Sodium tris EDTA
NTP	Nucleotide
PCR	Polymerase chain reaction
PK	Proteinase K
PKC	Protein kinase C
PMSF	Phenylmethylsulfonyl fluoride
RNA	Ribonucleic acid
RNase	Ribonuclease
SDS	Sodiumdodecylsulfate
TNT	Tris sodium tween
Tris	Tris[hydroxymethyl]-aminomethane
TSA	Tyramide signal amplification
WBS	Williams-Beuren Syndrome
X-Gal	5-bromo-4-chloro-3-indolyl- β -galactosidase

I. Introduction

The human and mouse genomes are now sequenced to an extent that had not been expected until the middle of this decade (Lander et al 2001, Mouse Genome Consortium 2002). This provides us with an approximate gene content which is 30000 for the human genome. As a result of this remarkable progress, the need to decipher gene function has become even more pressing. To advance understanding of gene function it is necessary to study gene expression at tissue and cellular levels. All the cells in our organism contain the same genome codified in the DNA, but not all cells are the same: a neuron for example, is very different from a blood cell or a muscle cell. This difference is due to the fact that in each type of cell a different genetic program is activated, turning on or off particular genes in various stages of embryonic development and adult-life. This phenomenon is described with an appropriate term: it is said that cells “express” different genes. This research has identified the expression-profile of the HC21 genes (see below), that is to say, a catalogue of all the genes of HC21, which are either on or off in different tissues during embryonic development. The outcome will eventually be a gene expression atlas, depicting where (in which tissue) and when (at what stage of development) the HC21 genes are expressed. The fact that a gene is expressed, at a particular stage of embryonic development, in a specific tissue or organ, may indicate that this gene has a function in that tissue (organ) or an important role in developmental regulation (Reeves 2002). The same is true even more so for pathological situations: if a gene is differentially expressed in a pathological compared to a healthy cell, this suggests that it might be implicated in the origins of the associated disease.

The cellular resolution of the non-radioactive RNA ISH technology presents the opportunity to visualize gene expression within specific cells in a tissue which therefore assists us in unraveling gene function.

1.1 Large scale expression analysis

To analyze the expression of genes by RNA ISH on a genome-wide scale, efficient technologies need to be used to determine and represent expression patterns of thousands of genes within an acceptable time scale. The objective here would be to scale-up the existing conventional gene-by-gene analysis approach to a genome-wide scale. This requires, high-throughput technologies that allow the determination (automated solvent delivery platforms) and visualization (automated microscopy) of the gene expression

patterns on tissue sections by RNA ISH at cellular resolution. The instrumentation used in this study was developed in our laboratory, where a catalyzed reporter deposition (CARD) based signal amplification technique, which improves the sensitivity of the non-radioactive ISH protocol, was adapted to the robotic platform (Bobrow et al. 1989). Although there have been previous large scale whole mount ISH studies which provide the spatiotemporal distribution of transcripts (Strachan et al. 1997, Neidhardt et al. 2000, Bulfone et al. 1998, Gawantka et al. 1998), this study is the first large-scale application of the automated non-radioactive ISH technique on tissue sections.

The technology required for automated high throughput ISH experiments on tissue sections and the machinery for automated high resolution microscopy that makes it possible to acquire images of a large number of specimens is outlined in Figure 1. For the efficient storage and retrieval of scientific data, web-based databases are excellent instruments. Therefore, a database has been developed that can provide all these features. The task of the database is to be able to import and store data (images and experimental metadata) and, when required, to allow retrieval of the data in an efficient manner (Visel et al. 2002).

Mouse embryonic day 14.5 was chosen because of the detailed expression information one can acquire from whole embryo sections. Unlike earlier stages, E14.5 mouse embryos show relatively well developed organs and tissues. The mouse is an excellent model organism both because it is easy to collect specimens and because it is a species with a well characterized genome suitable for the study of human biology (Bradley 2002, Smithies 1993). Being a mammal and having similar genetic and physiological systems, when compared to human, it only shows a difference in size and shape. Not only do the whole organ systems in the mouse show the same structure but also homeostatic equilibria, reproduction and behavioral parallels are observed (Bradley 2002). This causes many pathological conditions to arise in a similar manner in both species. Most human mutations and genetic diseases cause the same defects to show up in the mouse. On the other hand if the same phenotype is not observed when a defect is caused in the mouse orthologue gene, as is the case for the cystic fibrosis gene, the reason for this effect to not arise could give an answer on how to prevent the problem from occurring in humans. Another important reason favoring the mouse over other animals is that the laboratory mouse is tractable. One must not also forget that today the scientific community is experienced in manipulating the mouse genome. Today it is also possible to acquire ES cells of genes knocked-out via a gene trap approach from consortiums around the globe.

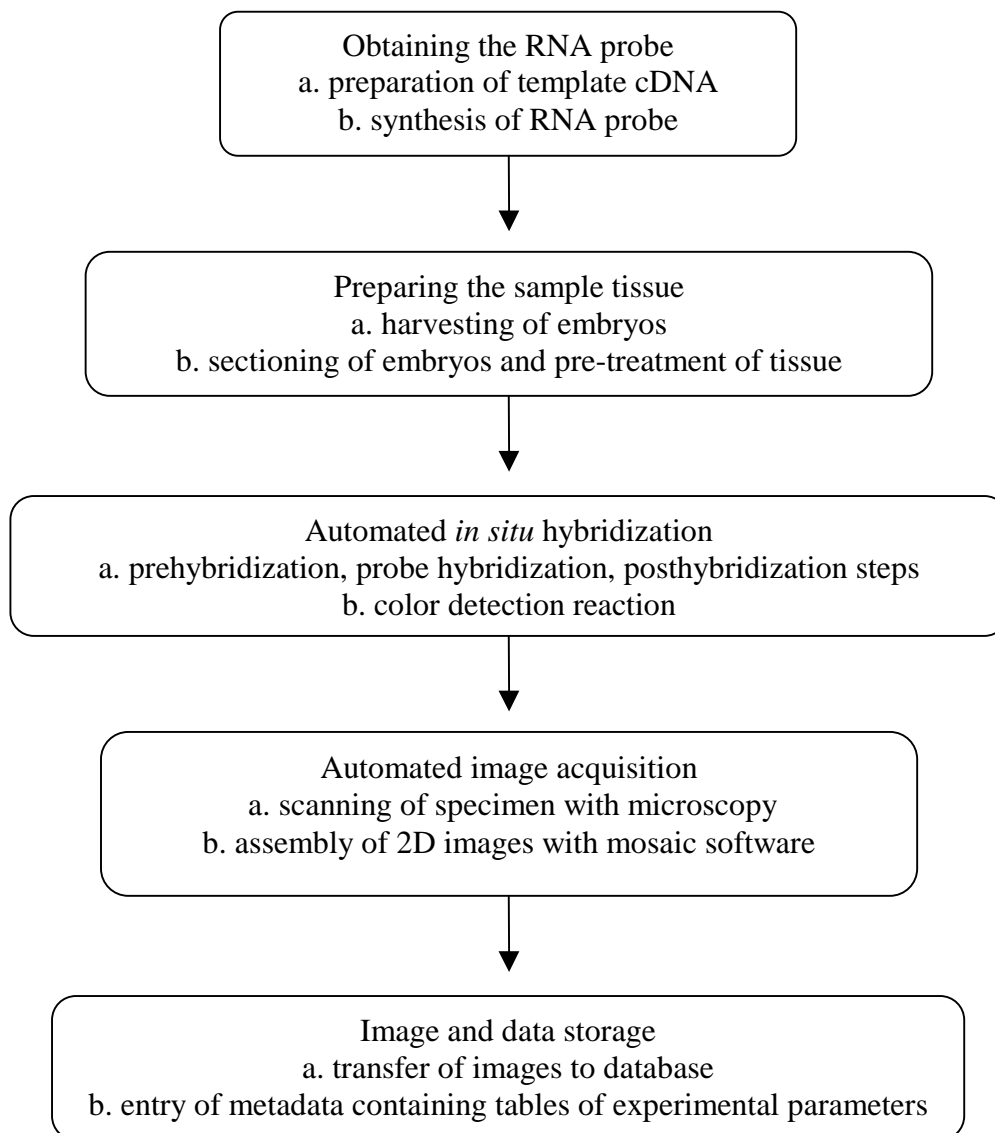


Figure 1. Flowchart representing the non-radioactive ISH procedure (GenePaint). The procedure has distinct elements, each of which can be automated to a considerable extent. Modified from Herzig et al 2001.

A systematic analysis of spatiotemporal gene expression patterns carried out in *Xenopus* during early development indicated that the identification of ‘synexpression groups’; genes with a shared complex expression pattern, can predict an association of genes to specific molecular pathways (Gawantka et al. 1998). Gawantka and coworkers have systematically analyzed 1765 randomly picked cDNAs by whole mount ISH.

Although we are able to utilize today -in principle- the whole mouse genome, when initializing this thesis project, it was necessary to concentrate on a well mapped region of

the human/mouse genome. Nearly the whole (99.7%) of human chromosome 21q had been sequenced by May 2000 and was publicly available, making it a very good object for our study (Hattori et al. 2000). Chromosome 21q represents ~1% of the human genome. The whole sequence of HC21 recently followed that of chromosome 22q (Dunham et al. 1999). Being correlated to an important disease like DS is one of the reasons why HC21 has played a leading role in the Human Genome project. A dense linkage map was first prepared for HC21 (McInnis et al. 1993), later a HC21 *NotI* restriction map was developed and the size of the long arm was estimated to be around 34 megabases based upon pulse-field gel electrophoresis studies using restriction fragments (Ichikawa et al. 1993). Today a 99.7% coverage of chromosome 21q has been achieved and 281,116bp have been sequenced from the short arm, where analysis of the base composition has revealed 178 confirmed genes and 36 predicted genes (Hattori et al. 2000, Annatorakis 2001, Davidson et al. 2001, Pletcher et al. 2001, Reymond et al. 2002a, Reymond et al. 2001). The mouse syntenic regions (conserved chromosome blocks containing orthologue genes which are consecutively ordered in an identical manner across species), which are segments of mouse chromosomes 10, 16 and 17 containing the mouse orthologues of the HC21 genes, harbor 170 orthologue genes (Figure 2). In other words the regions of homology that correspond to the human chromosome 21 are distributed on three different mouse chromosomes. 199 cDNA fragments representing 150 mouse orthologues (more than one fragment for some genes) were isolated for ISH experiments. This study demonstrates that it is now technologically feasible to determine the expression pattern of all the genes of the mouse genome.

1.2 Diseases related to chromosome 21

An extra copy of HC21 causes Down syndrome (DS) which affects 1 out of 700 live births and is the cause of severe mental retardation (Wiltshire et al. 1999). HC21 is the second smallest chromosome after human chromosome 22, containing approximately 200 genes, but its study is of enormous importance for research on DS and for a considerable number of other genetic diseases such as myopathy, anemia, blood platelet disorders, sclerosis, heart rhythm problems, one form of epilepsy, at least two forms of deafness and two types of cataracts (Hattori et al. 2000). The symptoms of DS are complex and involve various organs: the most frequent characteristics are mental retardation, low stature, short limbs, oblique palpebral fissures (upward slant to the eyes), protruding tongue, cardiac defects (present in 40% of the patients), a higher incidence of leukemia and defects of the

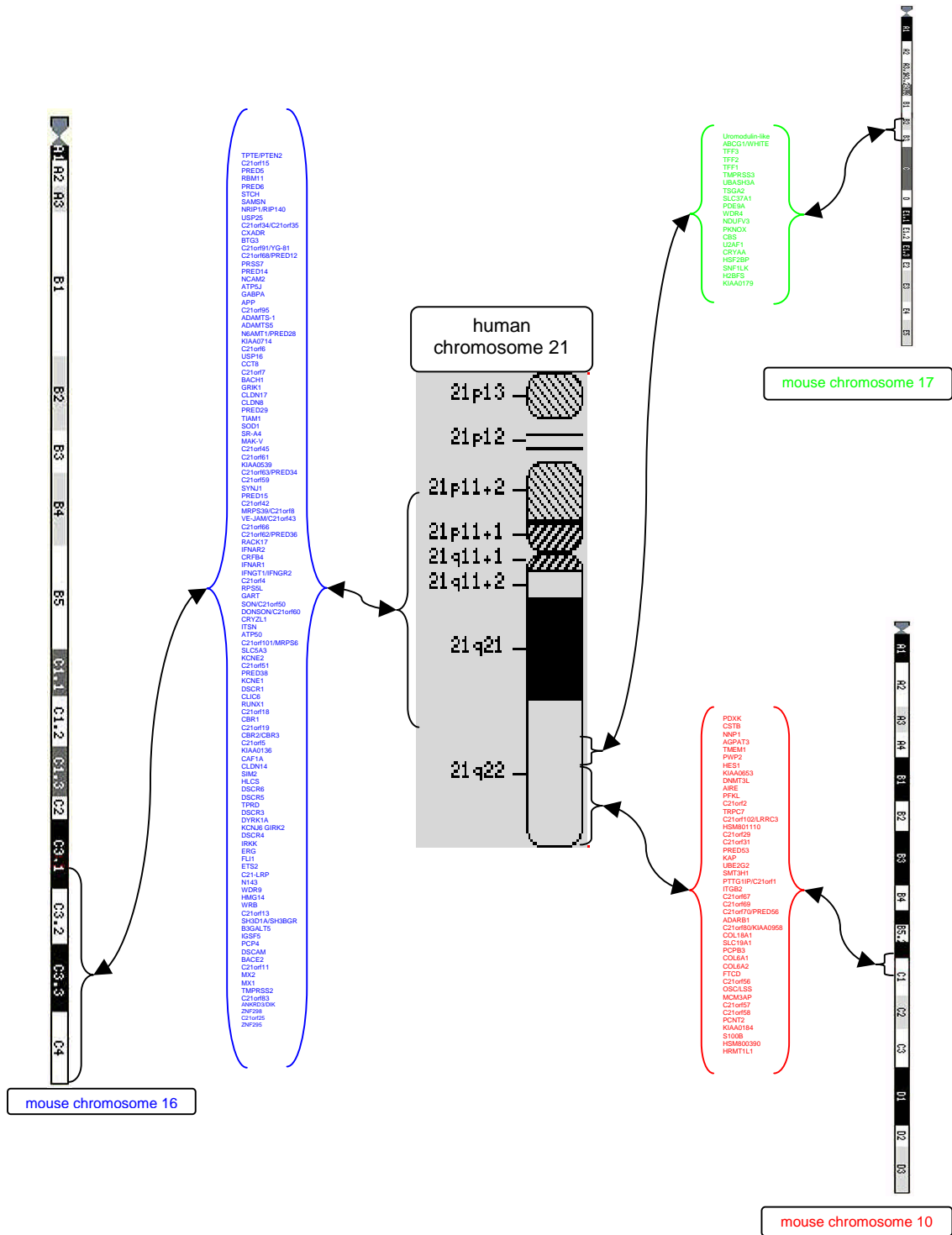


Figure 2. The synteny between human chromosome 21 and mouse chromosomes 16 (blue), 17 (green) and 10 (red) are depicted. All the known mouse orthologues of the HC21 genes are dispersed across three syntenic regions on three mouse chromosomes.

immune system (Epstein 1995). Furthermore, the amyloid beta precursor protein gene is localized on HC21, and this gene has been related to Alzheimer's disease. It is known that people with DS have early onset Alzheimer with higher frequency. Mutations in 14 known HC21 genes have been identified as the cause of monogenic disorders including one form of Alzheimer's disease (*APP*), amyotrophic lateral sclerosis (*SOD1*), autoimmune polyglandular disease (*AIRE*), homocystinuria (*CBS*) and progressive myoclonus epilepsy (*CSTB*). Loci for the following diseases which have been associated to a HC21 gene (due to mutations observed in patients) do not have their mouse orthologues pinpointed yet; recessive non-syndromic deafness (DFNB10 and DFNB8) (Bonne-Tamir et al. 1996, Veske et al. 1996), Usher's syndrome type IE (Chaib et al. 1997), Knobloch syndrome (*COL18A1*, Sertie et al. 1996), Bethlem myopathy (*COL6A1*, Jobsis et al. 1996), Ullrich disease (*COL6A2*), long QT6 (*KCNE2*) and holoprosencephaly type I (*HPEI*) (Estabrooks et al. 1990). The expression atlas of HC21 will give us a more precise insight on the role of these genes, which, to a large extent, is still unknown.

Partial duplication and deletions of other chromosomes exist, but only a few trisomies of whole chromosomes have been known to survive. The reason for their higher survival rate might be the lower density of genes present in HC21. Trisomies of other chromosomes are mostly embryonic lethal. Trisomies have been reported for chromosomes 13 and 18, but are observed at a much lower frequency compared to trisomy 21 (Beaudet et al. 1995). At present, no treatment to cure or prevent DS exists. The current care provided for the patients, though important, can only monitor and correct some of the symptoms such as those related with cardiac defects, and rehabilitation therapy which aims to support harmonic development and a good scholastic, social and professional integration.

Loss of heterozygosity was observed for specific regions of HC21 in several solid tumors (Sakata et al. 1997, Kohno et al. 1998, Ohgaki et al. 1998, Yamamoto et al. 1999, Ghadimi et al. 1999, Bockmuhl 1998) including cancers of the head, neck, breast, pancreas, mouth, stomach, esophagus, and lung. Loss of heterozygosity occurs when one allele of a gene is inactivated, in most cases through a mutation or chromosomal deletion, thus resulting in the loss of the functional allele. This usually indicates that there is at least one tumor suppressor gene in this region. Recently it has been shown that AML1-ETO can suppress transcription of the *p14^{arf}* tumor suppressor, through inactivation of the *AML1* (*RUNX1*; located on HC21) DNA binding site of the *p14^{arf}* promoter (Linggi et al. 2002). The decreased incidence of tumors in DS individuals indicates that an increased

dosage of HC21 genes may protect the cells of these individuals from transformation (Satge et al. 1998a, Hasle et al. 2000).

Due to the fact it is currently not feasible to carry out research on gestation or early gestation periods in humans, interest has turned to animal models. One approach was to create a segmental mouse trisomy model; this would be done by creating a segmental mouse trisomy for chromosome 16. Most of the HC21 genes are harbored on the mouse chromosome 16 syntenic region, which makes it the appropriate chromosome to be used in a mouse model representing the trisomy 21 disease state. Ts65Dn is a segmental mouse chromosome 16 trisomy that shows male sterility and degeneration of the cholinergic neurons of the basal forebrain (Reeves et al. 1995). A limitation of the Ts65Dn mouse is that the mouse chromosome 16 trisomy segment harbors more than just the HC21 orthologue genes. Another perspective would be to individually analyze genes by developing transgenic mouse models where the increased dosage of HC21 genes can be examined. Data from transgenic mice indicate that a subset of genes on HC21 may be involved in the phenotypes observed in DS patients (Kola and Herzog 1997).

In humans it is not known which and how many of the HC21 genes are important to cause the symptoms and signs of the syndrome. The expression atlas of HC21 is an important instrument for connecting the expression of these genes with the various symptoms of DS. It also gives us insight to which and how many genes are involved in the origins of this disease. For example, some of the genes of HC21 are expressed in particular regions of the heart during embryonic development; therefore these genes may have a role in congenital heart defects found in Down patients. In the same way, some genes that are expressed in the embryonic limbs like *Adams1* and *Erg*, might be involved in hand and feet anomalies typical of the syndrome. Finally, a set of HC21 genes are expressed in the brain, suggesting that they may play a role in mental retardation. Having studied E14.5 mice which correspond to human mid embryonic stages will allow this study to provide significant information on which HC21 genes could be important for the correct development of organs which especially have developmental defects in DS.

1.3 Relationships existing between expression patterns and the organization of the genome

The significance and importance of the expression pattern and the annotation atlas of the HC21 genes in relation to genetic disorders is obvious. The illnesses occurring through mishaps that affect the genes or gene clusters located on HC21 will be pinpointed,

eventually allowing us to prevent or treat these diseases. Different from any previous studies, an expression pattern analysis of a whole chromosome presents the opportunity to explore whether there might be a relationship existing between the organization of genes on a chromosome and the expression pattern of these genes. The reading of the genome requires that its individual components are correctly organized to permit the proper readout and hence a natural development of an organism. To achieve this, the genome needs to synchronize and influence its components through communication carried out across various levels of hierarchy. This could be done through the co-regulation of genes belonging to shared regulatory pathways, owing to shared non-coding sequence motifs that direct the binding of specific groups of shared transcription factors (Li et al. 1999). Several of these sequence motifs are similar to cis-regulatory elements found in the regulatory regions of other known developmentally regulated genes. Thus genes localized across the genome involved in the same genetic network would be regulated in harmony due to shared regulatory elements. The Hox family of genes, which play a crucial function in the patterning of the body, are another example where a non-random organization of the genes in the genome is observed. The Hox genes are clustered as a block of genes on one chromosome (Boncinelli et al. 1993). Analysis of the sequence of the 5' region of the Hox genes has allowed the identification of known sequences capable of interacting with transcription factors (Awgulewitsch et al. 1990). Therefore the consecutive placement of genes could allow their coordinated regulation through a regulatory element. This is due to the fact that the parallel transcription of these genes might be simultaneously required at a certain stage of development.

However, beside the known classical elements like promoters, classical enhancers, chromosomal insulators and matrix or scaffold attachment regions, there is the probability that the co-regulation of genes might be in response to Locus Control Regions (LCRs) (Li et al. 1997, Schreiber and Bernstein 2002). LCR confer high level, tissue specific, integration site independent, copy number dependent expression on linked transgenes (Pennacchio and Rubin 2001). Further studies are required to assess this mode of gene regulation, but attention needs to be drawn to the point that our expression data for HC21 provides the basis of such work. Hence through this study we not only wish to elucidate the pathogenesis of diseases but we expect to gain insight into novel mechanisms of gene control.

II. Materials and Methods

2.1 Materials

Table 1. The chemicals and kits were purchased from the following companies.

Product	Supplier
[$\alpha^{32}\text{P}$] dCTP	Amersham Pharmacia
Agarose	Fluka
Ampicillin	Sigma
Ampuwa sterile water	Fresenius
Anti-DIG-Fab-POD	Roche
Bacto-Agar	Difco, USA
Bacto-Yeast Extract	Difco
BCIP	Roche
Big dye terminator Mix	PE Applied Biosystems
Bisbenzimidazole (Hoechst)	Sigma
BMB blocking reagent	Roche
Bromphenolblue	Sigma
BSA, bovine serum albumin	Sigma
Calciumchloride	Biomol
Chloroform	Roth GmbH, 76185 Karlsruhe
Deckgläser, 24x55, No1, Superior	Marienfeld
Dextran sulfate	Sigma
Dnase I (RNase free)	Roche
DNTPs	Roche
DTT, dithiothreitol	Sigma
EDTA	Sigma
Ethanol	Roth GmbH, 76185 Karlsruhe
Ethidium bromide	Sigma
Ficoll, MW 400,000	Sigma
Formaldehyde	Electron Microscopy Sciences
Formamide	Calbiochem
Glutaraldehyde, 25% Lösung	Merck, 64271 Darmstadt
Glycerin	Gibco BRL
Glycin	Sigma
Hydrochloric Acid, HCl, 1N solution	Roth GmbH, 76185 Karlsruhe
Hydrogen Peroxide	Fluka
Hydro-Matrix	Micro-Tech-Lab
ISH Buffer	Ambion
Iodoacetamide	Sigma
IPTG	Biomol
Kaiser's glyceringelatine	Merck, 64271 Darmstadt
Kb-Ladder	NEBiolabs
KH_2PO_4	Merck, 64271 Darmstadt
Levamisole	Sigma

Materials and Methods

Liquid scintillation cocktail (Ready Flow III)	Beckman
Maleic Acid, disodium salt, anhydrous	Fluka
Methanol	Roth GmbH, 76185 Karlsruhe
Microscope slides	Menzel Gläser
Na ₂ HPO ₄	Riedel de Haen
NBT	Roche
NEN blocking reagent	NEN/Perkin-Elmer
NEN tyramide signal amplification kit	NEN/Perkin-Elmer
NeutrAvidin-AP	Pierce
Phenol	Roth GmbH, 76185 Karlsruhe
Plasmid-Miniprep-Kit	Qiagen, USA
Polyvinylpyrrolidone, PVP	Sigma
Potassium Chloride, KCl	Roth GmbH, 76185 Karlsruhe
Proteinase K	Roche
Restriction enzymes	NEBiolabs
Reverse Transcriptase	Invitrogen
RNase A	Sigma
RNase Inhibitor	Roche
RNaseZAP	Sigma
Salmon-Sperm-DNA	Sigma
SDS	Serva
Select Peptone	Gibco BRL
Sheep Serum	Chemicon International Inc.
Sodium Chloride, NaCl	Roth GmbH, 76185 Karlsruhe
Sodium Hydroxide, NaOH, 1N solution	Roth GmbH, 76185 Karlsruhe
T3-RNA-Polymerase	Roche
T4-DNA-Ligase	Gibco BRL
Taq-DNA-Polymerase	Qiagen, USA
Tissue-Tec (OCT)	Sakura
Triethanolamine	Fluka
Tris	Roth GmbH, 76185 Karlsruhe
Trisodium citrate dihydrate	Roth GmbH, 76185 Karlsruhe
Tween-20	Sigma
X-Gal	Biomol
Yeast tRNA	Roche

2.2 Methods

2.2.1 Isolation of nucleic acids

2.2.1.1 Mini preparation for plasmid DNA isolation

Minipreps were prepared by inoculating a single *E. coli* colony into 4mL of LB medium containing the appropriate antibiotics (ampicillin or kanamycin, 50µg/µL), in 15mL sterile culture tubes by using a sterile toothpick or loop and incubated overnight in the shaker (300rpm) for 16-18 hours at 37°C. After preparing a glycerol stock of each culture (200µL glycerol+ 800µL culture), the samples were centrifuged at 1400g for 15 min at 4°C. After removing the supernatant, the bacterial cell pellet was resuspended in 100µL of resuspension buffer P1 and transferred into Eppendorf tubes. Then, 150µL of lysis buffer P2 was added into the samples and mixed gently by inverting 4-6 times, followed by the addition of 150µL of neutralization buffer P3 and mixed immediately but gently by inverting 4-6 times. The cells were then centrifuged (25 min at 12,000g, at 4°C). After centrifugation, the protein and chromosomal DNA pellet was discarded and the supernatant containing plasmid DNA was transferred into a new Eppendorf tube. The plasmid DNA was precipitated by adding 1mL of 100% Ethanol (0.7 volumes) and centrifuged (25 min at 12,000g, at 4°C), the pellet was further washed with 1mL of 70% ethanol. The pellet was dried at room temperature and dissolved in 30µL water. "Qiagen Plasmid Midi Kit" (Qiagen, USA) buffers were used in this procedure.

2.2.1.2 Midi preparation of plasmid DNA isolation

In order to isolate a larger amount of plasmid DNA, the midi preparation DNA isolation method was carried out. For this purpose, the "Qiagen Plasmid Midi Kit" (Qiagen, USA) was used. 50µL of miniprep culture (as described in 2.2.2.1.) was used to inoculate 50mL LB medium containing the appropriate antibiotics (ampicillin or kanamycin, 50µg/µL) and incubated overnight in the shaker (300rpm) for 16-18 hours at 37°C. After preparing a glycerol stock of the culture (200µL glycerol+ 800µL culture), the bacterial suspension was centrifuged at 1400g for 15 min at 4°C, and the pellet was then resuspended in 4mL of Buffer P1. For the cell lysis, 4mL of Buffer P2 was added, then mixed gently but thoroughly by inverting 4-6 times, the suspension was incubated at room temperature for 5 min. 4mL of Buffer P3 was added to neutralize the suspension and mixed by inverting the tube. The suspension was centrifuged at 12,000g for 30 min and the supernatant applied to a previously equilibrated Qiagen-tip100 column (by applying 10mL of Buffer QBT for column equilibration) and left to run through the column resin by gravity flow. The column was washed twice with 10mL of Buffer QC. The DNA was eluted with 5mL of Buffer QF, and precipitated by adding 3.5mL (0.7 volumes) of isopropanol. Which

was then centrifuged immediately at 12,000g for 30 min at 4°C. The pellet was washed with 70% ethanol, air dried for 10 minutes at room temperature and dissolved in 50µL water. The DNA concentration was subsequently determined with a photometer, (see 2.2.2).

2.2.2 Determination of the nucleic acid concentration

The concentration of DNA or RNA was determined photometrically (8452A photometer, Hewlett Packard, Hamburg) by measuring absorption of the diluted sample, at 260nm and 320nm (control). The concentration was calculated according to the following formula:

$$C = (E_{260} - E_{320}) \times f \times c$$

C = concentration of the sample (µg/µL)

E₂₆₀ = absorption at 260nm

E₃₂₀ = absorption at 320nm

f = dilution factor

c = concentration_(standard)/absorption_(standard)

for double stranded DNA; c = 0.05µg/µL and for RNA; c = 0.04µg/µL

2.2.3 Restriction of DNA

For the specific digestion of plasmid DNA, type II restriction endonucleases were essential. These enzymes are capable of recognizing, binding and cleaving short defined nucleotide sequences in the target DNA. The restriction of the DNA was carried out in a defined buffer system, when multiple enzymes were used, a single buffer system such as "One-phor-all" (Amersham Pharmacia) was used. Digestion of DNA was carried out for 1- 2 hours at optimal temperature. One unit of enzyme is enough to cleave 1µg of DNA (having a single cleavage site) in 1 hour at 37°C. In a 20µL reaction, the following components were added:

1µg DNA
 2µL 10x reaction buffer
 1µL enzyme (10 units)
 xµL H₂O
20µL (total volume)

2.2.4 Ligation of DNA fragments

The ligation of foreign DNA fragments into vectors was carried out in the following reaction mix and the reaction was incubated overnight at 16°C. Although various ratios may be set up for a better chance of ligation, the vector to fragment ratio is usually 1:3.

30ng digested vector DNA
 90ng foreign DNA fragment
 2µL T4 DNA ligase enzyme (5U/µL)
 4µL 5x ligation buffer
 xµL H₂O
20µL (total volume)

2.2.5 Transformation of competent bacterial cells

A 50µL volume of competent cells (*E.coli* DH5α or TOP 10 cell) were mixed with 5-10µL of ligation reaction and placed on ice for 30min. The transformation reaction was then incubated at 42°C for 90 sec and placed on ice for 2 more minutes. In order to accelerate bacterial growth, 200µL of LB medium was added to the suspension which was incubated at 37°C for 45 min. After incubation, 100-150µL of the suspension was spread on the appropriate plate (Oja or kanamycin plate). The plate was incubated overnight at 37°C. The selection of a positive colony was carried out with the blue-white screening method. The insert containing colonies were white, while the others were blue (when the bluescript vector was used).

2.2.6 Preparation of LB medium and LB agar plates

2.2.6.1 LB medium

For 500mL LB medium, 5g tryptone, 2.5g yeast extract and 5g NaCl were dissolved in 450mL of deionized water. The pH of the solution was then adjusted to 7.2 with 5M NaOH, and the volume increased to 500mL with deionized water. The sample bottle was autoclaved for 25 min on liquid cycle. After autoclaving and cooling, the necessary antibiotics were added.

2.2.6.2 LB agar plate with kanamycin

For the preparation of 500mL LB agar, 5g tryptone, 2.5g yeast extract, 5g NaCl and 7.5g agar were dissolved in 450mL deionized water, then pH was adjusted to 7.2 and finally the volume was completed to 500mL by deionized water. The LB agar medium was autoclaved for 25 min, left to cool to approximately 55°C. After adding 1mL of IPTG (100mM) and 500µL of Kanamycin (50 mg/mL), the medium was poured into Petri dishes. The plates were stored at 4°C, if not used immediately.

2.2.6.3 Oja plate

500mL LB agar was autoclaved, then it was left to cool to approximately 55°C, the antibiotic, such as 500µL of Ampicilin (100mg/mL) and other components: 1mL of X-Gal (2%) and 1mL of IPTG (100mM) were added. The mixture was poured into Petri dishes and stored at 4°C.

2.2.7 Agarose gel electrophoresis

Non-denaturing gel electrophoresis was employed for the separation of restricted DNA. The concentration of the gels used, varied between 0.8% and 2% agarose content (w/v) in 0.5x TAE buffer. The TAE buffer was also used as electrophoresis buffer.

The agarose was dissolved in the buffer by boiling. The solution was slowly cooled, then 2.5µL of ethidium bromide (0.5µg/mL) was added and poured into the gel bed. Before loading samples, 0.2 volume of stop mix (loading buffer) was added and mixed. The samples were then loaded into the wells of the gel and electrophoresis was carried out at a steady voltage (50 or 100 volts).

2.2.8 Polymerase Chain Reaction (PCR)

A typical PCR reaction was conducted as follows. Amplification of target sequences was carried out by mixing the following:

5µL	10x PCR buffer
10µL	5x Enhancer
1µL	10mM dNTP (50x)
2µL	upstream primer (10pmol)
2µL	downstream primer (10pmol)
variable µL	DNA (50-250ng)
0.4µL	Taq DNA Polymerase (2U)
variable µL	H ₂ O
50µL	(total volume)

The reaction was carried out in an Eppendorf, thin walled PCR tube (200µL). The Eppendorf PCR cycler used for the reaction had a heated-lid (105°C) which prevented the reaction volume from changing due to evaporation. First the target DNA was denatured by heating the reactants at 95 °C for 2 min. This is followed by 35 cycles:

20sec - 94°C (denaturation), 20sec - 55°C (annealing), 1min - 72°C (extension) (1min per 1000 base-pairs)

The last elongation step was extended to 9 min. After finishing the PCR, 5µL of the reaction was mixed with loading buffer and run on a 1.5% agarose gel.

2.2.9 Non-radioactive dye terminator cycle sequencing

The non-radioactive sequencing was achieved with a Dye Terminator Cycle Sequencing Kit (Applied Biosystems) and the reaction products were analyzed with automatic

sequencing equipment, namely a 377 ABI Prism DNA Sequencer (Applied Biosystems). For the sequencing reaction, four different dye labeled dideoxy nucleotides were used. These, when exposed to an argon laser, emit fluorescence which could be detected and interpreted. The reaction progressed in a total volume of 10 μ L containing 1 μ g plasmid DNA or 100-200ng purified PCR products, 10pmol primer and 2 μ L reaction mix (contains dNTPs, dideoxy dye terminators and Taq DNA polymerase). Elongation and chain termination took place during the following program in a thermocycler: 5 minutes denaturing followed by 25 cycles of; 95°C 15 sec, denaturing; 55°C 15 sec, annealing; 70°C 4 min, elongation. After the sequencing reaction, the DNA was precipitated with 1/10 volume Na Acetate and 2.5 volumes 100% ethanol. After washing the pellet with 70% ethanol to remove the salt, the pellet was air dried and then dissolved in 3 μ L of loading buffer, denatured at 90°C for 2 min, and finally loaded onto the sequencing gel.

2.2.10 Reverse transcription

1-5 μ g total RNA was mixed with 1 μ L of oligo (dT)₁₆₋₁₈ primer (500 μ g/mL) in a total volume of 10 μ L. Synthesis of cDNA was carried out with the ThermoScript™ RT-PCR System, Invitrogen.

Mix1

Oligo(dT)-Primer	1 μ L
RNA	up to 9 μ L (1 μ g/ μ L)
Depc-treated water	up to 9 μ L

To avoid the possible problems with secondary structure of the RNA which might interfere with the synthesis, the mixture was heated to 70°C for 10 minutes, and then quickly chilled on ice.

Mix2 (for one probe)

5x cDNA Synthesis Buffer	4 μ L
0,1M DTT	1 μ L
RNaseOut™ (40U/ μ L)	1 μ L
Depc-treated water	1 μ L
10mM dNTP-Mix	2 μ L

Vortex and mix, mix1 and mix2 and after heating to 42°C for 2 minutes add 1 μ L ThermoScript™ RT (15U/ μ L). Incubate for 60 min at 42°C in a water bath. Stop the reaction by heating to 70°C for 15min. Add 1 μ L RNase H (10U/ μ L) and incubate 20min at 37°C. Place on ice.

2.2.11 Quantitative PCR and the determination of detection sensitivity

To evaluate our detection sensitivity we chose a gene showing low ubiquitous expression throughout the tissue. The *Dscr3* gene was a suitable candidate for this purpose. The outline of the experiments conducted are presented in Figure 3.

2.2.11.1 Preparing P7 mouse brain sample for mRNA extraction

The expression pattern for the *Dscr3* gene was observed to be ubiquitous in the near-rostral portion of a postnatal P7 mouse brain. The brain was dissected out of the skull and the piece of brain tissue corresponding to the region of interest was collected after the rest was removed with a razor. Three P7 brain specimens were collected in this manner from NMRI litter mates. One brain (specimen Y; 193mg) was embedded in OCT for later cryo-sectioning. The other specimens were reserved for DNA (specimen 1; 189mg) and RNA (specimen 2; 196 mg) extraction.

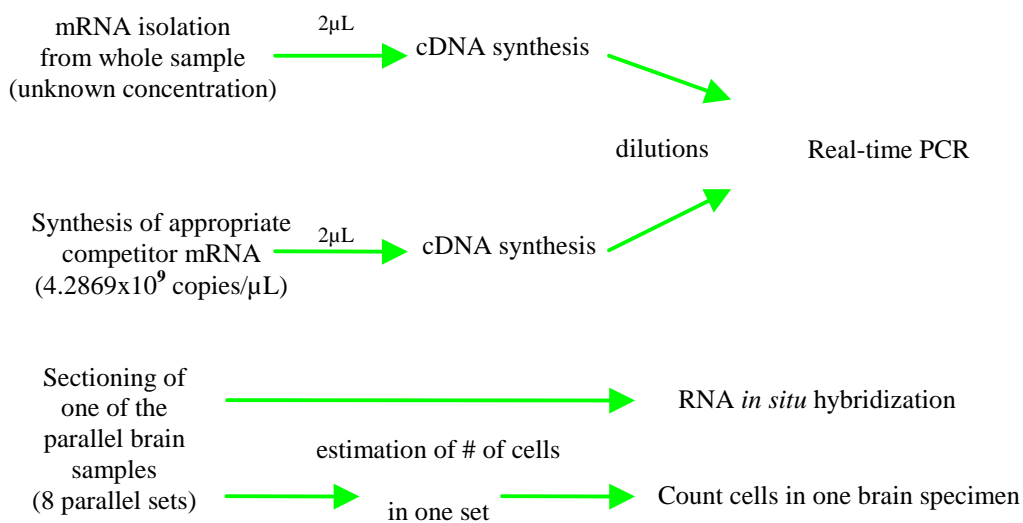


Figure 3. The diagram presents an outline of the experiments carried out to estimate the sensitivity of the instrumentation. The mRNA was isolated with Dynabeads as explained (2.2.11.2) and the standard was prepared using the Ambion competitor synthesis kit. The cell number was determined in one of the parallel sets after the approximate estimation of the cell radius.

2.2.11.2 RNA isolation from specimen

The mRNA was isolated from the 193mg brain tissue using the Dynabeads Oligo (dT)25 procedure as suggested by the manufacturer (DynaL Biotech). After initial isolation of the mRNA from the lysate, a second round of fresh Dynabeads were incubated with the

lysate to ensure all the mRNA was captured. The mRNA concentration was determined in a spectrophotometer and run on a gel to assess its quality.

2.2.11.3 Radioactive *in vitro* transcription

The Ambion Competitor Construction Kit was used to obtain high yields of *in vitro* transcription products in the 300-500 nucleotide range. As described by the manufacturer the following reaction cocktail was prepared:

1 µg template DNA in a volume not to exceed 5µL
 2 µL 10x Transcription buffer
 2 µL 5x (2.5mM ATP, GTP, 2' modified CTP and UTP) NTP solution
 2 µL T7 RNA polymerase (20U)
 <5 µL RNase-free dH₂O
 0.4µL [α -³²P] ATP (10mCi/mmol) (Amersham Pharmacia Biotech)
 20 µL total volume

The contents were mixed, and incubated at 37°C for 4 hours. 1µL was removed to a scintillation vial with appropriate cocktail. The removal of template DNA was carried out with DNase I (5U/µL) incubation as described by the manufacturer. Competitor was purified by running on a 5% acrylamide gel and extracted from the gel as described by the manufacturer. Finally the competitor was eluted in 15µL of nuclease free water and 1µL was transferred to a scintillation vial.

CPM readings were taken in a LKB Wallac 1209 Rackbeta Liquid Scintillation counter. The Beckman Ready Flow III Liquid Scintillation Cocktail was used for the procedure. CPM readings were taken for 60 seconds. Using the formula specified by manufacturer the copy number for the competitor was determined:

$$\frac{\text{CPM}/\mu\text{L of purified competitor RNA}}{\text{CPM}/\mu\text{L of transcription reaction}} \times \frac{3 \times 10^{14} \text{ molecules of ATP}/\mu\text{L}}{\# \text{ of A's in the competitor RNA}}$$

$$\frac{5578.1}{339452.2} \times \frac{3 \times 10^{14}}{115} = 4.2869 \times 10^{10} \text{ copies of competitor RNA}/\mu\text{L}$$

2.2.11.4 Determination of *Dscr3* mRNA concentration

The *Dscr3* gene was analyzed using the following PCR primer set: *Dscr3* (*Dscr3*F: AACTGCGCTATCACGCAGC; *Dscr3*R: TCCTGTTGCATTCCAGACG). The first strand cDNA was synthesized with oligo-dT primers (Thermoscript RT-PCR System,

Invitrogen). 1 μ L of mRNA (total elution volume of isolated mRNA: 30 μ L) was incubated with primers and dNTPs at 65°C for 5 minutes before cooling on ice. After adding 4 μ L of the appropriate buffer, 40U of RNase inhibitor and 2 μ L of 0.1 M DTT, the volume was adjusted to 19 μ L with RNase free water. The RT mix was incubated for 2 minutes at 42°C, and then 1 μ L Thermoscript reverse transcriptase was added (15U/ μ L) (Thermoscript RT-PCR System, Invitrogen). The RT reaction was incubated at 42°C for 60 minutes and then stopped by heating at 70°C for 15 minutes. The reaction was incubated with 1 μ L RNase H (10U/ μ L) (Thermoscript RT-PCR System, Invitrogen) for 20 minutes at 37°C.

Real-time PCR was carried out using 2 μ L of the cDNA. To analyze the *Dscr3* RNA a standard was generated. To construct the standard, the same primers used for Real-time PCR were used. To amplify the *Dscr3* specific fragment the following PCR conditions were used; denaturation for one cycle at 94°C for 2 minutes, 35 cycles of 94°C for 15 seconds, 56°C for 15 seconds, 72°C for 30 seconds and a final cycle at 72°C for 10 minutes. The PCR reaction was carried out in 25 μ L with final concentrations according to manufacturer's protocol (1 x PCR buffer, 1x Q-solution, 0.5U Taq polymerase, Taq PCR Core Kit, Qiagen) and sufficient dNTPs (0.5 μ L of 50x PCR Nucleotide Mix, Roche). The Taq polymerase was added after the 2 minute denaturation step. Amplified fragments were sub cloned into a vector pCRII-TOPO (Invitrogen). The orientation of the fragment was checked via sequencing. The pCRII-TOPO vector was used to prepare a DNA template, by PCR, that was sequenced and used to generate RNA by *in vitro* transcription according to the manufacturer's protocol (RT-PCR Competitor Construction Kit, Ambion). The RNA was purified, the concentration measured and the copy number was calculated. (see 2.2.11.3)

The Real-time PCR amplification was carried out in an iCycler iQ (Biorad) to determine the mRNA copy number for the *Dscr3* gene in the tissue sample. The PCR reaction (50 μ L) was set up according to the manufacturer's protocol (IQ SYBR Green Supermix, Biorad), with 1 μ L template (standard with a known copy number or cDNA). The reaction was carried out under the following condition in the iCycler: denaturation for one cycle at 95°C for 3 minutes, 45 cycles of 95°C for 30 seconds, gradient 53.5-56°C for 30 seconds, 72°C for 30 seconds where the real-time analysis was carried out at 72°C. The PCR products were checked by electrophoresis to confirm that a single PCR product was obtained. After a preliminary run using standards with 10 fold serial dilutions an

optimized run that was bracketing the unknown sample was done with 2 fold serial dilutions to generate the standard curve.

2.2.11.5 Isolation of DNA from Mouse Brain Sample

After the brain specimen was prepared (as described above), it was placed into an Eppendorf tube, then 700 μ L of lysis buffer and 35 μ L of Proteinase K (10mg/ml) were added. The sample was homogenized and incubated at 37°C for 1 hour. 350 μ L of phenol and 350 μ L of chloroform was added and sample was centrifuged at 12,000g for 10min, at 4°C. The upper phase was transferred to a new Eppendorf tube and 1 volume of chloroform was added. The sample was centrifuged and later the upper phase transferred to a fresh tube. 1/10 volume of Na Acetate (3M, pH 8) and 0.7 volume of isopropanol were added and the same centrifugation step was carried out. The pellet was washed with 1mL of 70% ethanol and then dried at room temperature. At the end, the pellet was dissolved in 100 μ L of 1x TE buffer.

2.2.12 cDNA's representing the mouse orthologues of the HC21 genes

We translated the putatively encoded proteins from the HC21 confirmed genes. These strings of residues were used to “tblastn” search the non-redundant database for orthologous murine cDNA sequences. We identified corresponding murine ESTs using the “blastn” program with the above isolated longest murine cDNAs as search queries. HC21 predicted genes and HC21 confirmed genes with no orthologous cDNA sequences in Genbank were compared directly to murine dbEST sequences to identify putative orthologous EST sequences. The availability of the Celera Genomics mouse genome sequence (<http://www.celera.com/>) allowed us to confirm that the identified ESTs/cDNAs were representative of the murine orthologues of the HC21 genes by *in silico* mapping. HC21 shows conserved synteny to mouse chromosome (MC) 16, 17 and 10 (<http://www.informatics.jax.org/>). As the human *ITSN* gene encodes a long and a short alternatively spliced transcript with ubiquitous and brain-specific expression (Guipponi et al. 2001), we collected and analyzed by ISH different murine ESTs specific for each form. In addition because *ERG* and its paralogue *Fli1* are extremely similar at the DNA level, we performed ISH with probes for both genes to assure that the *ERG* probe staining was specific. The different IMAGE and BMAP EST clones used in this study were obtained from Research Genetics (www.resgen.com). The National Institute of Aging provided the laboratory with the NIA 15K Mouse cDNA Clone Set and the Japanese National Institute of Infectious Diseases the NIID clones. To confirm the identity of the

cDNA, each clone was subjected to 5' and 3' end sequencing on an ABI377 or an ABI PRISM3100 (Applied Biosystems). As no murine ESTs were available for the *C21orf13*, *C21orf29*, *C21orf57*, *C21orf68*, *C21orf83*, *C21orf95*, *Cldn17*, *Erg*, *Pred53*, *Prss7*, *Slc5a3* and *Uromodulin-like* orthologues, we generated probes for these genes by RT-PCR from different normal murine tissues, cloned into a pCRII-TOPO vector (Invitrogen) and sequenced. Human and mouse comparative mapping of HC21, MC16, MC17 and MC10 was used to design specific primers in highly conserved regions. The evolutionary breakpoints lie between *ZNF295* and *UROMODULIN-like* (MC16/MC17 breakpoint) and inside the *PDXK* gene (MC17/MC10) (Davidson 2001). Consistently the ESTs/cDNAs representing orthologues of human genes mapping between the centromere and *ZNF295* map to MC16, the ones corresponding to the human genes mapping between *UROMODULIN-like* and *KIAA0179* map to MC17, and those analogous to the human genes mapping from *PDXK* to the telomere map to MC10, for three of the genes there is ambiguous information about the mouse orthology. A noteworthy exception is the *H2bfs* gene. *H2BFS*, a histone *H2B* family *S* member, maps to 21q22.3 just centromeric from *KIAA0179*, and another copy of this gene maps to 6p21. It is likely that the latter represents the ancestral copy from which the other copy arose through a duplication event, as the murine orthologous gene maps to MC13, in a region syntenic with 6p21. Only one gene was reported to map on the short arm of the acrocentric HC21, the *TPTE* gene (Hattori et al. 2000). The human *TPTE* gene is present in multiple copies in the human genome, but the active copies are thought to lie on 21p, 13 and 15q. Its murine orthologue was mapped to a region of MC8 that shows conserved synteny with human 13q14.2-q21. This region of the human genome contains a partial and highly diverged copy of *TPTE* that is likely to represent the ancestral copy from which the other copies of *TPTE* arose through duplication events (Reymond et al. 2001). The attempts to amplify large cDNAs fragments of the *Znf298* and *Slc37a1* were unsuccessful.

All the 199 probes for the HC21 orthologous mouse genes have been sequenced and along with various other metadata the images of expression patterns are publicly available at the website www.tigem.it/ch21exp/ and in the www.genepaint.org database. All known HC21 genes are presented in Table 2 in their chromosomal order.

Table 2. All verified human chromosome 21 genes

Gene	human gene accession #	mouse ortholog accession #	Clone	library	clone accession	chromosome	probe+
TPTE/PTEN2	NM_013315	AJ311311					
C21orf15	AY040090						
PRED5	AP001660						
RBM11	AP001660	BB633189	UIMBH21ape-h040UI	BMAP	BE648228	MC16	4G6
			clone 183		M.L.Y lab	MC16	4H12
PRED6	AP001660						
STCH	NM_006948	AK021006	1907602	IMAGE	AI226564	MC16	1H6
SAMSN1	NM_022136	NM_023380	3822366	IMAGE	BF660993	MC16	2B4
RIP140/NRIP1	NM_003489	NM_008735	408685	IMAGE	AI893344	MC16	1F5
USP25	NM_013396	NM_013918	3708890	IMAGE	BE569282	MC16	4C11
			3597996	IMAGE	BE533229	MC16	4C12
C21orf34/C21orf35	AP001666						
CXADR	NM_001338	NM_009988	H3135F11	NIA 15K	BG087114	MC16	1D11
BTG3	NM_006806	NM_009770	1515802	IMAGE	AW823883	MC16	1H2
C21orf91/YG-81	NM_017447	AK005829	1495644	IMAGE	AI664407	MC16	5A1
			4482610	IMAGE	BG242627	MC16	4G7
			H3025B10	NIA 15K	BG078224	MC16	4G5
C21orf68/PRED12	AK022689	AK014255	clone nb 6		RT-PCR	MC16	3G8
PRSS7	NM_002772	NM_008941	clone nb 9		RT-PCR	MC16	3G11
PRED14	AP001679						
NCAM2	NM_004540	NM_010954	UIMBH2.2aop-e120UI	BMAP	BE649879	MC16	4G12
			clone 130		M.L.Y lab	MC16	4H8
PRED15	AL163227						
C21orf42	NM_058184						
MRPS39/C21orf8/MRPL39	AK000458	AF239728	1514828	IMAGE	AW913227	MC16	1H1
JAM2/VE-JAM/C21orf43	NM_021219	NM_023277	1195543	IMAGE	AA690843	MC16	3C3
			UIMBH0ajqe050UI	BMAP	BE656219	MC16	3C2
ATP5J	NM_001685	NM_016755	H3115C02	NIA 15K	BG085594	MC16	1C11
			3512913	IMAGE	BG277264	MC16	4C1
GABPA	NM_002040	NM_008065	1282513	IMAGE	AA866905	MC16	1G7
APP	NM_000484	NM_007471	H3132G02	NIA 15K	BG086954	MC16	1D10
C21orf95/CYYR1	AY061853	AY061854	300 bp, clone nb 6		RT-PCR	MC16	4G4
ADAMTS1	NM_006988	NM_009621	H3034B07	NIA 15K	BG078965	MC16	1B5
ADAMTS5	NM_007038	NM_011782	569515	IMAGE	AA288689	MC16	3E7
N6AMT1/PRED28	NM_013240	AK013667	3971724	IMAGE	BF714407	MC16	2H2
KIAA0714/ZNF294	AB018257		1329026	IMAGE	AA929353	MC16	3D5
			1516272	IMAGE	AW824144	MC16	3D4
C21orf6	NM_016940	NM_016924	3484719	IMAGE	BE286595	MC16	2A11
USP16	NM_006447	NM_024258	H3152C04	NIA 15K	BG075823	MC16	4A1
CCT8	NM_006585	NM_009840	H3018A08	NIA 15K	BG077645	MC16	1A8
C21orf7	AY033900	AY033899	1153184	IMAGE	AA645026	MC16	3C1
BACH1	NM_001186	NM_007520	3599795	IMAGE	BE380232	MC16	3C7
GRIK1	NM_000830	X66118	UIMBH1ann-a030UI	BMAP	BE864009	MC16	1E11
CLDN17	NM_012131						
CLDN8	NM_012132	NM_018778	H3064H02	NIA 15K	BG081413	MC16	2H1
PRED29	AP001709		608371	IMAGE	AA162070	MC16	4H10

Materials and Methods

Gene	human gene accession #	mouse ortholog accession #	Clone	library	clone accession	chromosome	probe+
TIAM1	NM_003253	NM_009384	2609690	IMAGE	AW320557	MC16	3E3
			3412641	IMAGE	BE633001	MC16	3E4
SOD1	NM_000454	M35725	H3130B11	NIA 15K	BG074045	MC16	1D8
SR-A4/KIAA1172	AF023142	U49058 Rn3	779633	IMAGE	AA466932	MC16	3E2
HUNK/MAK-V	NM_014586	NM_015755	H3010H08	NIA 15K	BG077115	MC16	1A6
C21orf45	NM_018944	AK012533	3825587	IMAGE	BF715138	MC2	4E2
			3383880	IMAGE	BG145540	MC2	4E3
C21orf61							
KIAA0539/C21orf108	NM_014825		3976165	IMAGE	BG146293	MC16	4H3
			1092114	IMAGE	AA670781	MC16	4H7
C21orf63/PRED34	AF358258	AF358257	UIMAO0aby-g-120UI	BMAP	BE652074	MC16	1E12
C21orf59	NM_021254	AK003413	3654930	IMAGE	BF011432	MC16	4A11
			4191729	IMAGE	BF539818	MC16	4B7
SYNJ1	NM_003895	NM_053476	3470497	IMAGE	BF023103	MC16	3G10
C21orf66	AY033903	AY033907	10F7R-H		RT-PCR	MC16	2G3
			10F1R-B		RT-PCR	MC16	3E11
			10F9R-Q		RT-PCR	MC16	3E12
C21orf62/PRED36	NM_019596	AK016554	1264334	IMAGE		MC16	4F6
RACK17/OLIG2	XM_047941	AB038697	UIMBH0ait-d060UI	BMAP	AI854441	MC16	1F1
IFNAR2	NM_000874	NM_010509	3600669	IMAGE	BE381041	MC16	2F2
CRFB4	NM_000628	NM_008349	3994647	IMAGE	BF167015	MC16	2C1
IFNAR1	NM_000629	NM_010508	H3118F09	NIA 15K	BG085842	MC16	1C12
IFNGT1/IFNGR2	NM_005534	NM_008338	1531332	IMAGE	AW988781	MC16	1H4
C21orf4	NM_006134	BC004841	H3137H08	NIA 15K	BG087296	MC16	2G7
RPS5L	NM_001009 HC19	NM_009095	H3112G03	NIA 15K	BG072599	MC7	1C8
			H3112G04	NIA 15K	BG072599	MC7	1C9
GART	NM_000819	NM_010256	H3047B10	NIA 15K	BG079975	MC16	1B9
SON/C21orf50	AF380179	NM_019973	H3035B06	NIA 15K	BG065775	MC16	1B6
DONSON/C21orf60	NM_017613	AF193608	1314881	IMAGE	AA929739	MC16	2E2
CRYZL1	NM_005111	AK008846	H3123G05	NIA 15K	BG073533	MC16	1D6
			1531726	IMAGE	BE134049	MC16	4C8
ITSN	NM_003024	NM_010587	3156638	IMAGE	AW910367	MC16	2A4
ITSN long form	AF064244	NM_010587	1246139	IMAGE	AA823724	MC16	3C4
ITSN short form	AF064243	NM_010587	313401	IMAGE	AI414524	MC16	3C5
ITSN short form			419417	IMAGE	W88185	MC16	3C6
ATP50	XM_009742	AF254738	697411	IMAGE	AA230539	MC16	1F12
C21orf101/MRPS6	AY061855	AY061856	H3100H09	NIA 15K	BG084441	MC16	4E4
SLC5A3	NM_006933	NM_017391	1162 bp, clone nb 1		RT-PCR	MC16	4G3
KCNE2	NM_005136	AK008619	UIMBH1aly-e-030UI	BMAP	BE864913	MC16	4D12
C21orf51	AY033902	AY033901	3152735	IMAGE	BE553636	MC16	2E11
PRED38	AP001720						
KCNE1	NM_000219	NM_008424	1222885	IMAGE	AA667912	MC16	2D11
DSCR1	NM_004414	NM_019466	639212	IMAGE	AA200984	MC16	4E6
CLIC6	AF448438	AF448440	1511135	IMAGE	AW990844	MC16	1G12
RUNX1	NM_001754	NM_009821	4008335	IMAGE	BF135151	MC16	2C3
C21orf18	NM_017438	AY037804	717391	IMAGE	AI466884	MC16	3G4
			H3104D09	NIA 15K	BG071899	MC16	1C5

Materials and Methods

Gene	human gene accession #	mouse orthologue accession #	Clone	library	clone accession	chromosome	probe+
CBR1	NM_001757	NM_007620	3986613	IMAGE	BF120878	MC16	2B6
C21orf19	AF363446	AF363447	H3015D11	NIA 15K	BG077443	MC7	2G11
			3468606	IMAGE	BF682834	MC7	2E12
CBR2/CBR3	NM_001236	AK003232	1066729	IMAGE	AA611976	MC16	1G3
C21orf5	NM_005128	NM_026700	849213	IMAGE	AI481251	MC16	2D4
KIAA0136	D50926		3470826	IMAGE	BF023393	MC16	2F1
CAF1A	NM_005441	AK011243/BC013532	3981861	IMAGE	BF100487	MC16	2F7
			3995650	IMAGE	BF165815	MC16	2C2
CLDN14	NM_012130	NM_019500	1431432	IMAGE	AI042730	MC16	1G9
SIM2	NM_005069	NM_011377	3168352	IMAGE	BE457094	MC16	2A5
HLC5	NM_000411		455800	IMAGE	AA023801	MC16	3E9
DSCR6	NM_018962	AB063284	H3111B09	NIA 15K	BG085288	MC16	2G6
DSCR5	NM_016430	NM_019543	3470896	IMAGE	BF023450	MC16	2A10
TPRD/TTC3	NM_003316	NM_009441	H3004A06	NIA 15K	BG063181	MC16	1A2
DSCR3	NM_006052	NM_007834	H3007H01	NIA 15K	BG063488	MC16	1A4
			2780726	IMAGE	AW743140	MC16	1H12
DYRK1A	NM_001396	U58497	3988308	IMAGE	BF162516	MC16	2B8
KCNJ6 GIRK2	NM_002240	NM_010606	UIMBH3aul-h-110UI	BMAP	AW494400	MC16	2H8
DSCR4	NM_005867						
IRKK/KCNJ15	NM_002243	NM_019664	4234899	IMAGE	BF782321	MC16	3H10
ERG	M17254	AB073078	749 bp, clone nb 6		RT-PCR	MC16	4G1
ERG			821 bp, clone nb 9		RT-PCR	MC16	4G2
FLI (ERG paralogue)	NM_002017	NM_008026	808955	IMAGE	AA467723	MC9	1G1
ETS2	NM_005239	NM_011809	4021851	IMAGE	BF143562	MC16	2F11
C21-LRP/DSCR2	NM_003720	AJ238270	H3129G11	NIA 15K	BG086724	MC16	1D7
			3972244	IMAGE	BF722699	MC16	4C9
			3970373	IMAGE	BF722174	MC16	4C10
N143	AJ002572						
WDR9	AJ292465	AJ292467	3657672	IMAGE	BG100363	MC16	4A12
HMG14/HMG1	NM_004965	NM_008251	H3120G04	NIA 15K	BG086014	MC16	1D4
WRB	NM_004627		4021269	IMAGE	BF142081	MC16	2C8
C21orf13	AL163279		650 bp, clone nb 7		RT-PCR	MC16	5A5
SH3D1A/SH3BGR	NM_007341	NM_015825	586149	IMAGE	AA138480	MC16	1F10
B3GALT5	NM_033173	NM_033149	MNCb-0324	NIID	AU067673	MC16	3E10
IGSF5	AL163280		4917803	IMAGE	BG863468	MC16	4G8
			4236489	IMAGE	BF784177	MC16	4G9
			1513545	IMAGE	AW990468	MC16	4H6
PCP4	NM_006198	NM_008791	UIMAH0acw-a110UI	BMAP	AI839872	MC16	3D8
DSCAM	NM_001389	NM_031174	614553	IMAGE	AA170935	MC16	4D1
BACE2	NM_012105	NM_019517	313341	IMAGE	AI605601	MC16	2C9
C21orf11	AJ409094	AF360358	1546160	IMAGE	BE135671	MC16	2E8
MX2	NM_002463	NM_013606	3374009	IMAGE	BE629286	MC16	4H4
MX1	NM_002462	NM_010846	4020700	IMAGE	BF143741	MC16	2C7
TMPRSS2	NM_005656	NM_015775	4017848	IMAGE	BF168078	MC16	2C6
			3979313	IMAGE	BF102443	MC16	2F5
C21orf83	AY063456	NM_023663	850 bp, clone nb 3		RT-PCR	MC16	5A3
ANKRD3/DIK/PKK	NM_020639	NM_023663	3979359	IMAGE	BF102436	MC16	2B5

Materials and Methods

Gene	human gene accession #	mouse orthologue accession #	Clone	library	clone accession	chromosome	probe+
ZNF298	AP001745						
C21orf25	AL163290		4017776	IMAGE	BF168443	MC16	2C5
ZNF295	NM_020727		3156533	IMAGE	AW910268	MC16	2A3
UROMODULIN-like	AP001745		668 bp, clone nb 2		RT-PCR	MC17	5A4
ABCG1/WHITE	NM_004915	NM_009593	3988239	IMAGE	BF159563	MC17	2B7
TFF3	NM_003226	NM_011575	1166710	IMAGE	AA710846	MC17	2D9
			1545811	IMAGE	BE136788	MC17	1H5
TFF2	NM_005423	NM_009363	438574	IMAGE	AI894032	MC17	1F7
TFF1	NM_003225	NM_009362	493047	IMAGE	AI893383	MC17	2D1
TMPRSS3	NM_024022	AJ300738	Lung 1.1		RT-PCR	MC17	3E5
UBASH3A	NM_018961	AK020009	3375785	IMAGE	BE627812	MC17	2A7
TSGA2	AK057315	NM_025290	H3075A12	NIA 15K	BG082335	MC17	1B11
SLC37A1	NM_018964						
PDE9A	NM_002606	NM_008804	3326827	IMAGE	BE456255	MC17	2A6
WDR4	NM_018669	NM_021322	4009256	IMAGE	BF138291	MC17	2F8
			671843	IMAGE	AA242309	MC17	1F3
NDUFV3	NM_021075	BC013640	3465040	IMAGE	BE687712	MC17	2A8
PKNOX	NM_004571	NM_016670	3415880	IMAGE	BE687087	MC17	3G6
CBS	XM_033010	BC013472	H3079G04	NIA 15K	BG082744	MC17	3C11
			H3079G12	NIA 15K	BG082752	MC17	3C12
U2AF1	NM_006758	AK012849	H3086D11	NIA 15K	BG070353	MC17	1B12
CRYAA	NM_000394	J00376	355457	IMAGE	W48180	MC17 htgs	2C11
			420571	IMAGE	W90910	MC17 htgs	1F6
HSF2BP	NM_007031	AK016553	H3125B05	NIA 15K	BG086370	MC17	4A3
SNF1LK	AP001751	NM_010831	1260296	IMAGE	AA855274	MC17	3D11
H2BFS	NM_017445	AK011516	3465513	IMAGE	BE688921	MC13	2A9
			1531132	IMAGE	AW988475	MC13	2E7
KIAA0179	XM_035973	AK011155	H3020A05	NIA 15K	BG064464	MC17	1A9
PDXK	NM_003681	NM_031769	H3013B08	NIA 15K	BG077257	MC10	1A7
CSTB	NM_000100	NM_007793	421131	IMAGE	W91030	MC10	3D2
NNP1	NM_003683	NM_010925	H3158F11	NIA 15K	BG088875	MC10	1E9
AGPAT3	NM_020132	NM_053014	H3158D09	NIA 15K	BG088853	MC10	1E8
TMEM1	NM_003274	AC009295	H3108A09	NIA 15K	BG085049	MC10	2H6
PWP2	NM_005049	AK018333	4015754	IMAGE	BF143119	MC10	2F9
C21orf33/HES1	NM_004649	NM_008235	H3139H01	NIA 15K	BG074828	MC10	1D12
KIAA0653	AB014553		H3103C07	NIA 15K	BG084637	MC10	1C4
DNMT3L	NM_013369	AJ404467	H3094C02	NIA 15K	BG084050	MC10	1C2
			1295738	IMAGE	AA919800	MC10	4C3
AIRE	NM_000383	NM_009646	1282549	IMAGE	AI510675	MC10	2G4
			1265557	IMAGE	AI552580	MC10	2G5
PFKL	NM_002626	NM_008826	H3025D11	NIA 15K	BG064930	MC10	1A12
C21orf2	NM_004928	BC010330	H3158A06	NIA 15K	BG088819	MC10	2G9
TRPC7	NM_003307	AJ344343	1195195	IMAGE	AA718075	MC10	4E9
C21orf102/LRRC3	AY061857	AY061858	693205	IMAGE	AI466988	MC10	4E11
HSM801110	XM_036086						
C21orf29	AP001754		857 bp, clone nb 4		RT-PCR	MC10	5A2
C21orf34	AP001754						

Materials and Methods

Gene	human gene accession #	mouse orthologue accession #	Clone	library	clone accession	chromosome	probe+
PRED53	AP001754		clone nb 6		RT-PCR	MC10	5A6
KAP	AL163300	NM_010670	1197995	IMAGE	AA727102	MC10	2D10
UBE2G2	NM_003343	NM_019803	4021048	IMAGE	BF143429	MC10	4B4
			4038176	IMAGE	BF181354	MC10	4B5
SMT3H1	NM_006936	NM_019929	H3023B06	NIA 15K	BG078058	MC10	1A11
C21orf1	NM_004339		H3044B12	NIA 15K	BG079708	MC10	2G10
ITGB2	NM_000211	NM_008404	3593958	IMAGE	BE377951	MC10	2B1
C21orf67	AF380178						
C21orf69	AY035381						
C21orf70/PRED56	AF391113	AF391115	3515288	IMAGE	BG370650	MC10	3G2
ADARB1	NM_015833	AF403109	976454	IMAGE	AA619858	MC10	3H7
C21orf80/KIAA0958	NM_015227	AK009301	3991098	IMAGE	BF163502	MC10	3H9
			2136258	IMAGE	AI956321	MC10	1H7
COL18A1	NM_030582	NM_009929	H3111D11	NIA 15K	BG085302	MC10	1C6
SLC19A1	NM_003056	NM_031196	1399806	IMAGE	AI115777	MC10	1G8
PCPB3	NM_020528	NM_021568	918000	IMAGE	AI596471	MC10	2D5
COL6A1	NM_001848	NM_009933	H3151F07	NIA 15K	BG088322	MC10	1E4
COL6A2	NM_001849	X62332	H3152G04	NIA 15K	BG088405	MC10	1E5
FTCD	NM_006657	BC010813	1277595	IMAGE	AA880297	MC10	2D12
C21orf56	AP001759		3664026	IMAGE	BF319576	MC10	4G10
			515785	IMAGE	AA066026	MC10	4G11
			514002	IMAGE	AA062323	MC10	4H5
OSC/LSS	NM_002340	U31352 Rn	H3037A02	NIA 15K	BG065929	MC10	1B7
MCM3AP	NM_003906	NM_019434	H3031B12	NIA 15K	BG078720	MC10	1B4
C21orf57	AY040873		clone nb 4		RT-PCR	MC10	5A8
C21orf58	AY039243						
PCNT	NM_006031	NM_008787	H3078B07	NIA 15K	BG069605	MC10	4E12
KIAA0184	D80006	AK019853	3497934	IMAGE	BE303319	MC10	2A12
S100B	NM_006272		1380986	IMAGE	AI035528	MC10	4A8
HSM800390	AL050065						
HRMT1L1	NM_001535	AF169620	574888	IMAGE	AA120755	MC10	1F8

Color code for table;

Not done.
No murine probes.
No murine orthologue.

2.2.13 RNA detection by automated *in situ* hybridization: instrumentation for high throughput gene expression analysis

RNA ISH has been used since the mid 1980's. The principle to this method is simple, where tissue or sections of tissues are probed with tagged synthetic DNA or tagged anti-sense RNA. The RNA detection by ISH procedure for this study is a multi step procedure consisting of steps that have been extensively optimized throughout the years. The technique consists of the following consecutive main steps:

- preparation of sample and fixation of tissue
- pretreatment of sample that has been mounted onto microscope slides
- hybridization of the probes
- post hybridization washes to remove non-specifically bound probe and detection of the hapten-labeled probe
- microscopic analysis of the sample, image acquisition and image archive

All the glassware used in this procedure have been baked at 180°C overnight and all solutions used in the steps prior to post hybridization washes are DEPC treated to prevent RNA degradation, where later solutions are only prepared with ultra-pure water. The current study has been carried out with 14.5 day old embryos but can be applied to other developmental stages and tissues.

After the plug is observed, a NMRI mouse is sacrificed on the afternoon of the 14th day. The day the plug is observed is presumed to be day 1. The embedding is done into optimal freezing medium (OCT) as explained below. The chambers used for this purpose have a copper plate bottom which is placed onto an aluminum surface that is cooled down to -70°C (Herzig 2001). The embryo is placed into the freezing chamber, orientated and then placed onto the aluminum surface which cools the chamber through the copper base. Consequently the chamber gradually freezes from the bottom, preventing the shock freezing of the tissue, which might be the cause of damage due to pressure accumulating at the core, resulting from expansion from the surface inwards. This could be the result of rapidly freezing samples in liquid nitrogen. Orienting the embryos is very important for obtaining stereotaxic sections, with a consistent or near consistent line of plane. In Figure 4 Nissl stained E14.5 embryos are shown, where you can see that the main organs of the body are represented. Only organs like the pancreas and spleen are still primordial.

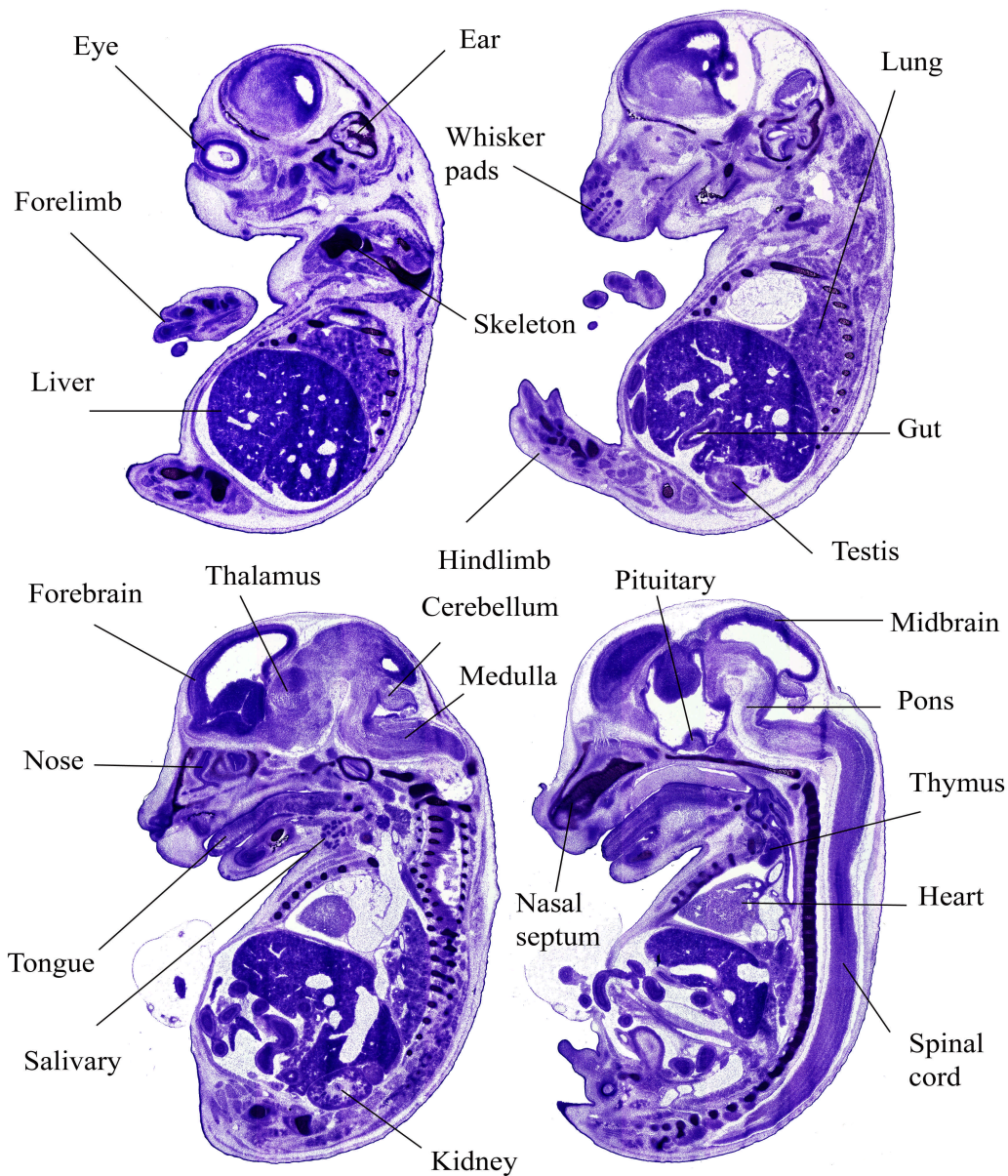


Figure 4. Nissl stained E14.5 embryos. In the above figure it can be seen that the major organs of the body have developed. Especially those of importance when looking for genes involved in Down's syndrome, like the brain, heart, limbs, gut and the thymus. This stage corresponds to mid embryonic human periods, giving it critical importance in our search for genes of developmental importance.

a. Preparation of the sample and fixation of tissue:

1. Freshly isolated embryos of the appropriate stage were dipped into ice-cold phosphate buffered saline (PBS, 0.137M NaCl 0.0027M KCl 0.01M Na₂HPO₄ 0.002M KH₂PO₄ pH 7.4) to wash away the blood. They were then transferred to ice-cold OCT and kept there for 5 minutes. All the steps are carried out on ice.
2. The embryo was transferred to an OCT embedding chamber, where they were frozen slowly. This was carried out with a freezing chamber specially designed for this

- purpose, which has a copper plate bottom that was in contact with an aluminum surface at -70°C .
3. The embryo was sectioned ($20\mu\text{M}$ thick) with a Leica CM3050S cryostat (Germany) with a sample temperature of -12°C and chamber temperature of -17°C . Sections were placed in appropriate positions on a Super Frost Plus microscopy slide (Germany). (see Figure 5 Cryostat)
 4. The tissue on the slides was fixed for 20 minutes with 4% paraformaldehyde (PFA) (EMS, USA) in PBS solution. Then the slides were washed in PBS twice to remove the fixative.
 5. The tissue was acetylated for 5 minutes (2x) with acetylation solution (0.25% acetic anhydride, 0,1M triethanolamine, pH 8.0, Sigma, Fluka respectively).
 6. The tissue on the slides was refixed for 20 minutes with 4% Para formaldehyde (EMS, USA) PBS solution. The slides were then washed for 5 minutes in PBS to remove the fixative, transfer to 0,9% NaCl for 5 minutes and dehydrate the cells in a graded series of ethanol (30%, 50%, 70%, 80%, 95% and 2x 100%) 2 minutes each step at room temperature. (Automated) (Figure 5 Autostainer XL)
 7. The slides were air dried. Slides are storable for up to 3 months at -80°C in air tight dessicated chambers.



Figure 5. On the left you can see the Leica CM3050S cryostat, the Leica autostainer XL can be seen on the right.

A liquid handling system retrieves and delivers buffers and reagents from receptacles placed next to the racks. Figure 6 depicts a Tecan Genesis platform adapted to the ISH protocol. The software controlling the platform is Gemini V3.2. Using this equipment with two racks (2 x 48 hybridization chambers), pre-hybridization, hybridization, post-hybridization and the hapten labeled probe detection steps were carried out automatically

with little human intervention or supervision. One run took approximately one day, where a daily throughput yielded as many as 380 sections.

Slides carrying the tissue sections were integrated into a flow through chamber (Figure 6). This device constitutes a small flow-through chamber with a 400 μ L reservoir on top, an 80 μ m thick chamber housing the tissue (volume; 120 μ L) and an exit for solutions on the bottom. Solutions filled into the well enter the narrow hybridization chamber by gravity and remain in place due to capillary forces until displaced by a fresh solution filled into the well. Eight such platforms are arranged into a row and six rows constitute a 48 position rack (Figure 6). In this setup the slide of the hybridization chamber directly contacts the surface of the platform permitting efficient thermal transfer.

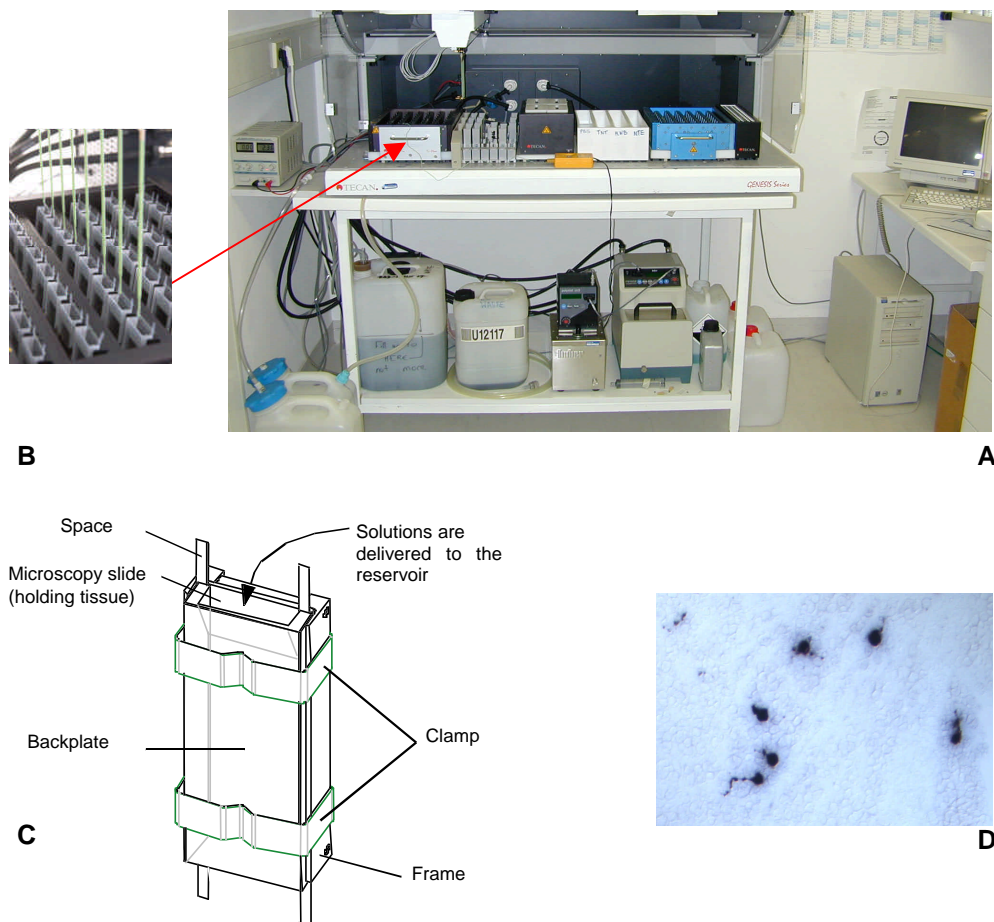


Figure 6. A Tecan Genesis platform modified to carry out automated high-throughput *in-situ* hybridization. (A) Genesis RSP 150 platform, (B) one of the two racks containing 48 places for hybridization chambers, eight such platforms are arranged into a row and six rows constitute a 48 position rack, (C) hybridization chamber with a 400 μ L reservoir on top, an 80 μ m thick chamber housing the tissue (volume: 120 μ L) and an exit for solutions at the bottom and (D) signal observed on tissue sections at higher magnification (*Npy*).

b. Pretreatment of the sample that has been mounted onto microscope slides (Automated): Unless otherwise stated all the following pretreatment steps were at room temperature. Allow slides to thaw before opening air tight chambers.

1. The slides mounted into slide chamber were treated, 6 times for 5 minutes with 300 μ L of MeOH solution (Roth) containing 0.6% hydrogen peroxide (Fluka). The slides were then washed 8 times with 300 μ L of PBS with 0.05% Tween20.
2. Then the slides were treated 2 times for 4 minutes with 300 μ L of 0.2N HCl with 0.05% Tween20. Next the slides were washed 8 times with 300 μ L of PBS with 0.05% Tween20.
3. The slides were treated 2 times for 7 minutes with 300 μ L of Proteinase K in buffer (with appropriate concentration corresponding to the stage of tissue to be treated) (50mM TRIS, 5mM EDTA, pH 8.0) with 0.05% Tween20. The slides were then washed 8 times with 300 μ L of PBS with 0.05% Tween20. Mild protease treatments have a positive effect on detection efficiencies and should be optimized depending on tissue type and concentration for each application.
4. Then the slides were treated 2 times for 10 minutes with 300 μ L of 4% PFA in PBS solution with 0.05% Tween20. The slides are then washed 8 times with 300 μ L of PBS with 0.05% Tween20. This step is very important for the retention of RNA and the preservation of cellular morphology, but its is critical to not over do this step to make sure too much cross linking does not prevent RNA accessibility.
5. Finally the probes were pre-hybridized with ISH Buffer (Ambion, Cat. # B8807G) for 30 minutes at 60°C.

c. Hybridization of the probes (Automated):

The probes used for transcript detection vary according to the variability of the ORF of the genes to be studied and to various other criteria that assure probe specificity. As a rule of thumb they are 1000bp probes from the 3 prime UTR. One reaction cocktail provides sufficient probe to be dissolved into the Ambion ISH Buffer (300 μ L per slide, final concentration of 100ng/mL). The probes were allowed to hybridize at 60°C over-night.

Procedure for obtaining the *in vitro* transcribed hapten labeled RNA Probe: The immunological detection procedure is described for the detection of a digoxigenin labeled RNA probe designed for the detection of specific gene transcripts. Standard labeling methods were used, varied as explained below depending upon the DNA source

that would provide the template for the *in vitro* transcription of the hapten labeled antisense RNA probe. The digoxigenin was detected with an anti-digoxigenin antibody coupled to peroxidase, which initiated the tyramide signal amplification step. As a result of the signal amplification, tyramide bound biotins are crosslinked to electron rich structures, like proteins and nucleotides in the vicinity of the specific transcript. Which then are detected by a neutravidin antibody coupled alkaline phosphatase. The alkaline phosphatase is what creates the precipitate out of the chromogenic substrates. With the help of an amplification step based on enzyme catalyzed reporter deposition we believe to have increased our sensitivity 1000 fold (Adams 1992). Antisense RNA probes with a length of 1000 nucleotides provide best results. It is important to ensure the probes are not too long, as it may be difficult for longer probes to penetrate the sample, although this is a greater problem with whole mount RNA ISH procedures, it is less of a problem while using sections. Even probes of 2 to 3kb in length have given us good results. One must take care to prevent cross reactivity against other homologous genes. Longer probes could also require a higher stringency washing temperature, so the T_m must be checked carefully. If the gene specific templates have arrived as a DNA template directly go to step 2. Otherwise the following steps are taken;

1. Primer design and reverse transcription: After the gene specific sequence, which will be the template for the later *in vitro* transcription reaction, is determined, 20-24 base long primers bordering this sequence region should be designed. Keep in mind that they fulfill the conditions required for a good primer. Then the sequence of the T7 or SP6 universal primers, with 3 extra bases (GCG) to the 5' end of the forward primer and the 3' end of the reverse primer are attached to allow the polymerase to bind better. (Important; please take care to attach the T7 primer to the forward and the SP6 to the reverse primer). The annealing temperature depends on the sequence of the gene specific part of the primer because the T7 (GCG-TAATACGACTCACTATAGGG) or Sp6 (GCG-ATTTAGGTGACACTATAG) sequences play no role in binding to the mRNA. Therefore, paste only the gene specific part of the primer in any program to determine the annealing temperature (either local or on the web). To determine the best annealing temperature, run an analytical (gradient) PCR with 6 tubes in a range of 20 degrees (PCR efficiency may change according to various cyclers). For the first PCR (using the RT product as template), 20 μ L reactions for 35 cycles should be carried out and 5 μ L of this reaction used for electrophoretic analysis. Do not forget the T7 and Sp6 or T3 primers (depending on the vector) used need to have the extra 3 bases (GCG) for better RNA polymerase binding. If some of the PCR's show the desired band, those PCR's are pooled

and purified with Qiagen's PCR purification kit (or equivalent). If there are other bands present, all left over samples are mixed and run on an agarose gel, the desired band is excised and the DNA purified with a gel extraction kit. 250ng are used for sequencing thus verifying the PCR product. If the sequence that will be used as a template for RNA transcription has already been sub cloned into a vector, 50ng of this DNA is taken and used for the reamplification process. The PCR is carried out at 60°C with a gradient up to 65°C with 3 tubes. 20µL reactions for 35 cycles should be carried out and 5µL used for electrophoretic analysis (Figure 7A). The sequences of all the probes are available as metadata on genepaint.org.

Table 3. Typical gradient program used in cycler:

Nr.	Temp (°C)	Time (min)	Cycles	Gradient
1	94	2.00	1	
2	94	0.20	34	10°C
3	60	0.20		
4	72	1.10		
5	72	9.00	1	
	4	Pause		

Table 4. PCR reaction

H ₂ O	6.9-7.9µL
10x Buffer	2µL
Q (enhancer)	4µL
dNTPs (2 mM, Roche)	2µL
Primer for (5 pmol/µL)	2µL
Primer rev (5 pmol/µL)	2µL
Template (RT-product)	1-2µL
Enzyme Taq-Polymerase (Qiagen)	0.1µL

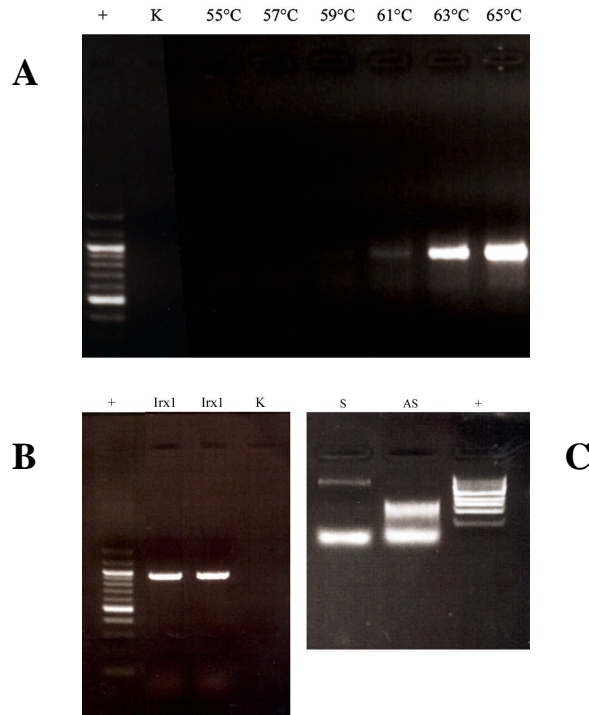


Figure 7. The first PCR, second PCR and reverse transcription (A) Gradient PCR showing band of interest around 1KB, “+” symbolizing the DNA standard. The first gradient PCR (using the RT product as template) carried out as a 20 μ L. The desired PCR products are the bands at the higher temperature of the gradient. (B) Second PCR later used as template for the *in vitro* RNA transcription. (C) Sense and antisense RNA obtained from the *in vitro* RNA transcription of the DNA template.

2. Template preparation: In order to prepare template for the RNA probe, a new 100 μ L PCR is carried out using the purified 1st PCR product as template (described above). If the probe has already arrived as a DNA template this can be treated as the 1st PCR product. Either the same primers as for the first PCR (now the T_m will be much higher because there will be a 100% match) or primers designed for sequencing can be used. The first alternative is preferable because on one side, primers designed for sequencing are shorter and therefore have lower T_m , and on the other, they lack the extra bases at the 5'end which are added to enhance the stability of the polymerase binding. Of course the first approach can be used for probes that have already arrived as DNA template. From the product of this preparative PCR, 5 μ L are analyzed on an agarose gel and the rest is purified with Qiagen's PCR purification kit and finally eluted with 25 μ L of buffer (10mM tris, pH8.5). 3 μ L are used to measure the concentration and determine DNA purity. Typically yields are in the range of 200-400ng/ μ L. See Figure 7B.

Table 5. Second PCR reaction

H ₂ O	~37.5µL
10x Buffer	10µL
Q (enhancer)	20µL
dNTPs (2 mM, Roche)	10µL
Primer for (5 pmol/µL)	10µL
Primer rev (5 pmol/µL)	10µL
Template (RT-product)	~2µL (20-50 ng)
Enzyme Taq-Polymerase (Qiagen)	0,5µL

3. *In vitro* RNA transcription: Currently the Strategene RNA transcription kit is used. 1µg of template DNA that was obtained as explained above is used for this procedure. The following are added up to a total volume of 20µL;

Table 6. *In vitro* RNA transcription reaction

H ₂ O (depc)		Add up to 20µL
5 x transcription buffer		4µL
rATP	10 mM	2µL
rCTP	10 mM	2µL
rGTP	10 mM	2µL
rUTP	10 mM	1,3µL
DIG-UTP	250 nmol	0,7µL
DTT	0,75 M	2µL
RNasin	40 U/µL	1µL
RNA Polymerase (T3 or T7 or SP6)	50 U/µL	0,5µL
Template	1 µg	~

DEPC water is added to obtain a volume of 20µL. The relevant RNA polymerase was purchased as follows; T3 from Strategene, T7 and SP6 from NEB. The reaction mix was incubated at 37°C for 4 hours and then the following were added to stop the reaction and remove the DNA template:

Table 7. DNase treatment of *in vitro* RNA transcription reaction

H ₂ O (depc)	16,4µL
MgCl ₂ (0,3 M)	1,6µL
DNaseI (10 U/µL)	2,0µL

The reaction is incubated for 30 minutes at 37°C after DNase I addition. Then was precipitated the mix with 72µL of NH₄ Acetate (4M) and 470µL of absolute ethanol. RNase free solutions were used. After adding the solutions store at -80°C for 30 minutes and centrifuge 10000g at 4°C. The salt was washed away with ice-cold 70% ethanol. Then the precipitate was eluted in 25µL of DEPC water and 2µL is checked on a gel, while 2µL was used to determine concentration in a photometer. After checking the integrity of the RNA probe, the end volume was adjusted to 50ng/µL for each probe and

aliquot them into 50 μ L volumes and store them at -80°C. Prior to usage we dilute the probe to a concentration of 10ng/ μ L and store it at -20°C for a maximum of 2 months after which it is discarded. See Figure 7C.

4. Radioactive *in vitro* RNA transcription: 1 μ g of template DNA is used which is obtained as explained above. The following are added up to a total volume of 20 μ L;

Table 8. *In vitro* RNA transcription reaction

H ₂ O (depc)		Add to 20 μ L
5 x transcription buffer		4 μ L
rATP	10 mM	0,5 μ L
rCTP	10 mM	0,5 μ L
rGTP	10 mM	0,5 μ L
S ³⁵ rUTP (Amersham)	10mCi/mL	5 μ L
DTT	0,75 M	1 μ L
RNasin	40 U/ μ L	1 μ L
RNA Polymerase (T3 or T7 or SP6)	50 U/ μ L	0,5 μ L
Template	1 μ g	~

DEPC water is added to obtain a volume of 20 μ L. The reaction mix is incubated at 37°C for 4 hours and then the following are added to stop the reaction and remove the DNA;

Table 9. DNase treatment of *in vitro* RNA transcription reaction

H ₂ O (depc)	16,4 μ L
MgCl ₂ (0,3 M)	1,6 μ L
DNaseI (10 U/ μ L)	2,0 μ L

The reaction is incubated for 30 minutes at 37°C after DNase I addition. The RNA was precipitated with 72 μ L of NH₄ Acetate (4M) and 470 μ L of absolute ethanol. RNase free solutions were used. After adding the solutions store at -80°C for 30 minutes and centrifuge 10000g at 4°C. The salt was washed away with ice-cold 70% ethanol. Then the precipitate was eluted in 25 μ L of DEPC water and 1 μ L was removed to a scintillation vial with the appropriate cocktail. CPM measurements were taken in a LKB Wallac 1209 Rackbeta Liquid Scintillation counter. The Beckman Ready Flow III Liquid Scintillation Cocktail was used for the procedure. Prior to usage the probe was diluted to a concentration of 50000CPM/ μ L. 200 μ L was added per slide in the standard hybmix. The radioactive RNA ISHs were removed from the robot after the stringency washes. Rinsed in water and dehydrated in a series of ethanol. Air dried and stored in boxes until further processing. The sequence and other metadata related to the probes used are available at genepaint.org (eg. *Irx1*, *Dtx1*, *Colbal*, *Btg3*, *Nurr1*, *Itsn*, *Npy*, *Clic6* and *Sh3bgr*).

d. Post hybridization washes to remove non-specifically bound probe and detection of the hapten-labeled probe (Automated):

Post hybridization washes are used to remove the non-specifically bound and unbound probes from the tissues, thus decreasing the background and preventing nonspecific signal from occurring, all the following post hybridization steps are at 65°C:

1. Wash 5 times for 6 minutes with 300µL wash solution 1 (5x SSC) (1x SSC, 0.3M NaCl 0.03M sodium citrate trisodium citrate dehydrate pH 7.0).
2. Wash 5 times for 6 minutes with 300µL of wash solution 2 (2x SSC, 50% formamide).
3. Wash 5 times for 6 minutes with 300µL of wash solution 3 (1x SSC, 50% formamide).
4. Wash 5 times for 6 minutes with 300µL of wash solution 4 (0.1x SSC).

The recommended dilution of the antibodies to detect the hapten-label within the probe (e.g., digoxigenin), indicated for the reconstitution volume (amounts are values suggested by the supplier).

1. Wash 3 times with 250µL of NTE (0.05% Tween 20, 5mM EDTA, 10mM Tris, 500mM NaCl, pH 7,6)
2. Incubate 2 times at room temperature, for 5 minutes each, by applying 250µL of, 20mM iodoacetamide diluted in NTE . Iodoacetamide decreases the background by removing the activity of various endogenous enzymes
3. Wash 3 times with 250µL of NTE
4. Incubate 2 times at room temperature, for 45 minutes each, by applying 250µL of heat-inactivated 4% sheep-serum (Chemicon International) in TNT solution (0.05% Tween 20 in 100mM Tris 150mM NaCl, pH 7,6) (filter before use, 0,45µm). This solution is to decrease the background by preventing non-specific binding of the antibody.
5. Wash 8 times with 250µL of TNT solution
6. Incubate 2 times at room temperature, for 30 minutes each, by applying 250µL of TNB blocking buffer (0.05% Tween 20 in 100mM Tris 150mM NaCl, 0,5% blocking reagent (PerkinElmer Lifesciences) pH 7,6) (filter before use, 0,45µm). This solution also decreases the non-specific binding of the antibody.
7. Incubate 2 times at room temperature, for 45 minutes each, by applying 250µL of Roche's Anti-Dig-POD diluted in TNB blocking buffer (1:500 diluted in TNB, filter through 0,45µm pore filter before adding antibody)
8. Incubate 8 times at room temperature, for 5 minutes each, by applying 250µL of TNT

9. Incubate 2 times at room temperature, for 15 minutes each, by applying 250 μ L of Tyramide-Biotin diluted with NEN amplification diluent buffer for use with tyramide signal amplification (TSA) (PerkinElmer Lifesciences)
10. Wash 8 times with 250 μ L of maleate wash buffer (MWB) (100mM Maleate, 150mM NaCl, 0.05% Tween 20, pH 7,5)
11. Incubate 2 times at room temperature, for 30 minutes each, by applying 250 μ L of Neutravidin-alkaline phosphatase conjugate (Pierce) diluted in 1% blocking reagent (Roche) containing MWB (1:600 dilution in 1% blocking reagent containing MWB, filter through 0,45 μ m pore filter before adding antibody)
12. Wash 8 times with 250 μ L of maleate wash buffer (MWB) (100mM Maleate, 150mM NaCl, 0.05% Tween 20, pH 7,5)
13. Wash 2 times with 250 μ L of TNT solution
14. Incubate 3 times at room temperature, for 20 minutes each, by applying 250 μ L of TMN solution (0.1M Tris, 0.05M MgCl₂, 0.1M NaCl, 0.05% Tween, pH 9,5) containing BCIP (1:250) NBT (1:200) (Roche) containing levamisol (0.5mg/mL) (Sigma). This step maybe extended depending on signal strength. Levamisol is for the inhibition of various mammalian alkaline phosphatases (Van Belle, 1972).
15. Wash 4 times with 250 μ L water (0.05% Tween)
16. Wash with 250 μ L TNT twice
17. Fix for 20 minutes with 4% PFA, 0,5% glutaraldehyde (Sigma) containing PBS (with 0.05% Tween) by applying 250 μ L twice
18. Wash 4 times with 250 μ L PBS (with 0.05% Tween)
19. Wash 4 times with 250 μ L water (with 0.05% Tween)
20. Let the slides air dry overnight and cover slip them the next day with aqueous cover slipping medium (Hydromatrix). Cover slipping is also carried out in a robot (Figure 8A). After cover slipping the extra cover slipping medium is cleaned. The slides are labeled and ready for microscopic analysis.

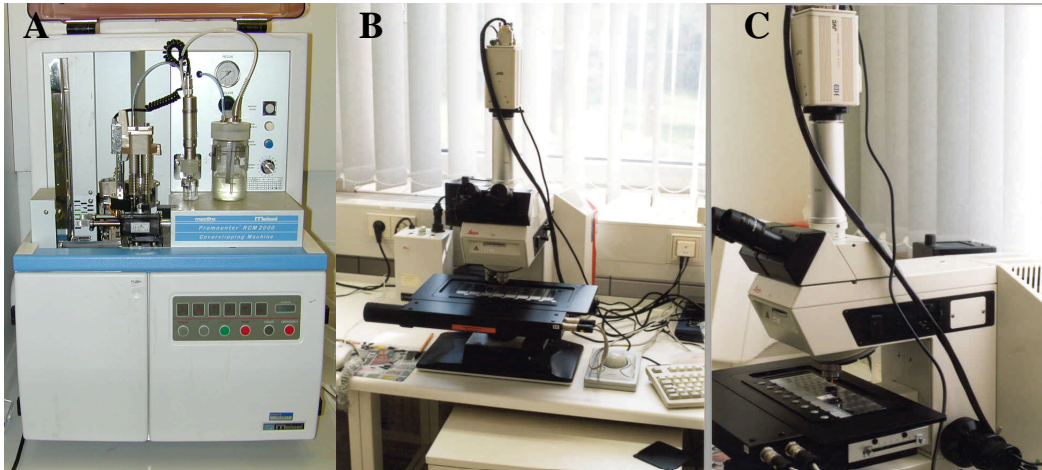


Figure 8. (A) Cover slipping robot (B-C) microscope with motorized stage and camera

e. Microscopic analysis of the sample, image acquisition and the image archive:

The signal of a non-radioactive ISH is observed as a dark blue signal that is imaged digitally in a bright field microscope at a magnification adjusted to single cell resolution. The massive amount of data generated by the ISH robot requires effective and automated image data acquisition. The fundamental issue to consider in image data collection is that of resolution. Although it is possible to increase the resolution where subcellular components are detectable, this is not necessary while developing a transcriptome atlas (Carson et al. 2002). Data acquisition will be managed with a compound microscope equipped with a scanning stage that accurately translocates the specimen in front of the objective (Figure 8). Because the whole embryo section can not be captured at the required magnification as a single image that fits into the object field of a compound microscope, multiple images collected through the help of a motorized stage are stitched together to produce a mosaic representing the entire section. Individual images are stored as bitmap files.

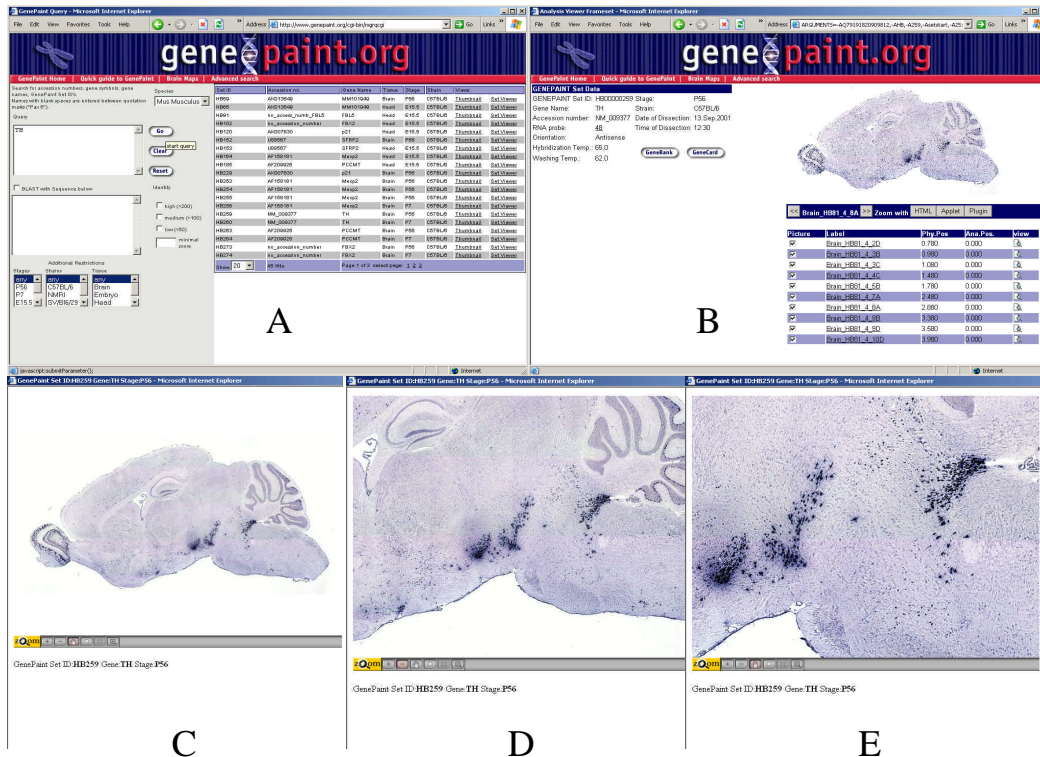


Figure 9. Main features of the Genepaint.org query and results page (A), the section-set viewer (B), incremental magnifications obtained with the virtual microscope demonstrating the capabilities of the zoom viewer (C-E). The applet allows zooming into the mouse adult brain specimen thus providing a detailed view of *tyrosine hydroxylase* expression at cellular resolution within the dopaminergic neurons of the *substantia nigra*. All the HC21 orthologue mouse genes are publicly available at genepaint.org.

Together with the website at www.tigem.it/ch21exp the images are located on an archive at the following web site genepaint.org (see Figure 9). Biological data is accessible on databases located on the internet (Discala et al. 1999, Ringwald et al. 2000). Web-based databases are an excellent means to efficiently retrieve and disseminate scientific data. Large quantities of data from a variety of specimens require a database, which is a powerful tool able to track data acquisition and storage. A database has been developed that can provide all these features. The task of the database is to be able to import and process data, allowing it to be later retrieved when required (Visel et al. 2002).

III. Results

3.1 Hapten labeled non-radioactive RNA *in situ* hybridization

It is a widely held view that hapten-labeled RNA ISH ("non-radioactive ISH") protocols are not as sensitive as radioactive RNA ISH methods. This would limit the usefulness of non-radioactive ISH. However, methods have been developed that enhance sensitivity. Specifically, tyramide signal amplification (TSA) allows up to 1000 fold amplification in the signal strength (Adams 1992). This figure is an estimate based on immunohistochemical assays on tissue sections. As TSA requires significantly more steps than standard ISH protocols, TSA is not frequently used. In the advent of automated ISH developed in our laboratory, this limitation can be overcome. A frequently held view of TSA-based ISH is that it has high background. Once more automation of the ISH process has allowed us to increase the number of washing and blocking steps resulting in a low background and hence a favorable signal to noise ratio. In this section we exemplify the difference of signal strength in the absence and presence of TSA amplification, illustrate the low background of our method, compare radioactive and TSA based non-radioactive ISH and provide a quantification of the sensitivity of the TSA method as executed in the automated ISH system.

3.1.1. Effect of Tyramide Signal Amplification and low background

To illustrate the potent effect of TSA amplification, we subjected E14.5 mouse embryo sections to ISH, probing mouse *Pcp4* mRNA in the presence or absence of a TSA step (Figure 10A). There is a remarkable difference in signal strength. While the TSA treated specimen shows a regional pattern of *Pcp4* transcripts in the brain, such signal is virtually absent in the specimen not processed with TSA. This is particularly clear in the enlarged view of the midbrain shown in Figure 10A. Also note the favorable signal-to-noise ratio in these images. To further explore this particular point, we examined gene expression in knock-out mice. Sections obtained from null mutants of *Ear2* (Figure 10B) and *Prkcb* (Figure 10C) were compared with sections obtained from their wild type littermates. In the case of *Ear2*, a weakly expressed gene, distinct signal is seen in many regions of the embryo (Figure 10B). In contrast, an equivalent section of an *Ear2* mutant embryo is virtually blank (Figure 10B). When high-power images of the olfactory epithelium are compared, this dramatic difference is still seen. We extended this analysis to another knock-out mouse in which the beta isoform of PKC was deleted (*Prkcb*). As can be seen in Figure 10C there was strong signal in many regions of the adult wild type brain but the

section of the mutant brain was once more blank. The high-powered images shown in Figure 10C emphasize this point.

The initial 1300 nucleotide probe used for examining the *Prkcb* null mutant sections contained a ~700bp stretch that had ~90% nucleic acid sequence identity with other PKC isoforms. As can be seen in Figure 11A this probe results in a detectable signal especially in the hippocampus, even in KO mice. This signal obviously represents cross-hybridization with other PKC isoforms and vanishes when a *Prkcb* specific probe was used (Figure 11B).

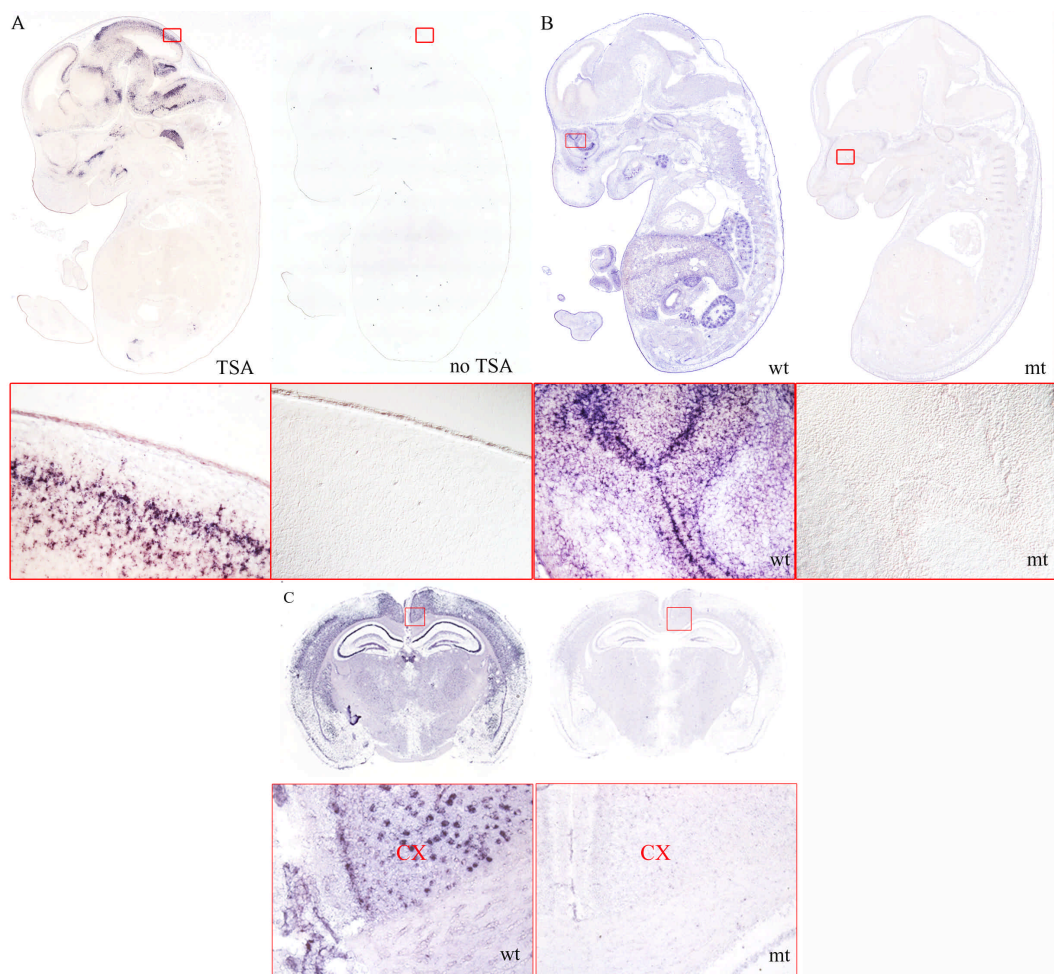


Figure 10. Sensitivity gained from TSA and low background. (A) Using the *Pcp4* gene antisense probe RNA ISH was performed including the TSA step (left) and excluding the TSA step (right). (B) *Ear2* mutant (mt) vs. *Ear2* wild type (wt). (C) *Prkcb* mutant vs. *Prkcb* wild type. The knockout specimens were hybridized with antisense probe of the appropriate gene; specifically the *Prkcbeta* was hybridized with adult tissue specimens sectioned coronally while the *Ear2* specimens were hybridized with E14.5 sagittal sections. It can clearly be seen that the wt specimens on the left show strong staining while the mt specimens on the right are nearly blank. Tissue in red boxes in the low power view are shown below each specimen at high power. Cortex (Cx).

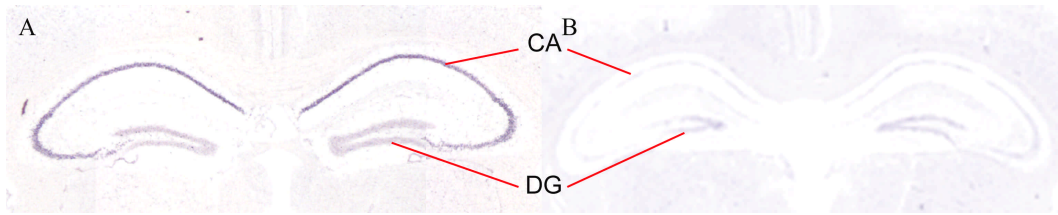


Figure 11. Cross reactivity due to non specific probe. *Prkcb* expression in brains using either a 1300bp (A) or a 700bp (B) probe. The longer 1300bp probe has homology to *Prkca* and thus the signal in the cornu ammonis (CA) stems from cross-hybridization as can be seen above (Dentate gyrus DG). Although the null mutant tissue was expected to look blank when hybridized with the *Prkcb* probe, the image on the left side shows this was not the case. Whereas on the right image, when the homologous region was removed from the probe, the cross-reactivity occurring in the hippocampus had disappeared.

3.1.2 Comparison of radioactive and non-radioactive ISH

To show that the hapten labeled RNA ISH protocol used in this study was within the range of sensitivity that is attributed to conventional (^{35}S -UTP) radioactive RNA ISH methods, we directly compared radioactive procedures with non-radioactive TSA enhanced ISH. Importantly, both types of data were from the same specimens and both were generated on the ISH robot. This comparison was done for a panel of genes that included *Clic6*, *Col6a1*, *Sh3bgr*, *Btg3*, *Irx1*, *Itsn*, *Npy*, *Nurr1* which showed a broad spectrum of expression patterns and levels (Figures 12, 13). Generally there was excellent agreement between the results obtained by the two methods. *Clic6* was locally expressed in the choroid plexus which is obvious with both methods. *Col6a1* shows expression in various parts of the heart. Note the strong signal in the mitral valve, apparent in both types of preparation. *Sh3bgr* (*Sh31da*) shows uniform expression in cardiac smooth muscle. *Btg3* transcripts were in a few cells migrating within the midbrain, bordering the cerebellum. Blue-stained cells are well visible in the non-radioactive specimen as are the silver grains resulting from ^{35}S ISH. *Irx1* is expressed within the tegmentum in a graded fashion as is apparent with both methods. Thus even gradients of gene expression can be revealed by the non-radioactive procedure. The same point was made from *Itsn* that not only shows very specific expression in the cortical plate but also a quantitative difference between cortex and thalamus. *Npy* was expressed in isolated cells within the inferior colliculus and these cells were seen in great detail with the non-radioactive ISH but less clearly with the autoradiograph. *Nurr1* was expressed in the midbrain hindbrain junction, most probably within the dopaminergic neuron precursors. This emerges from both types of data.

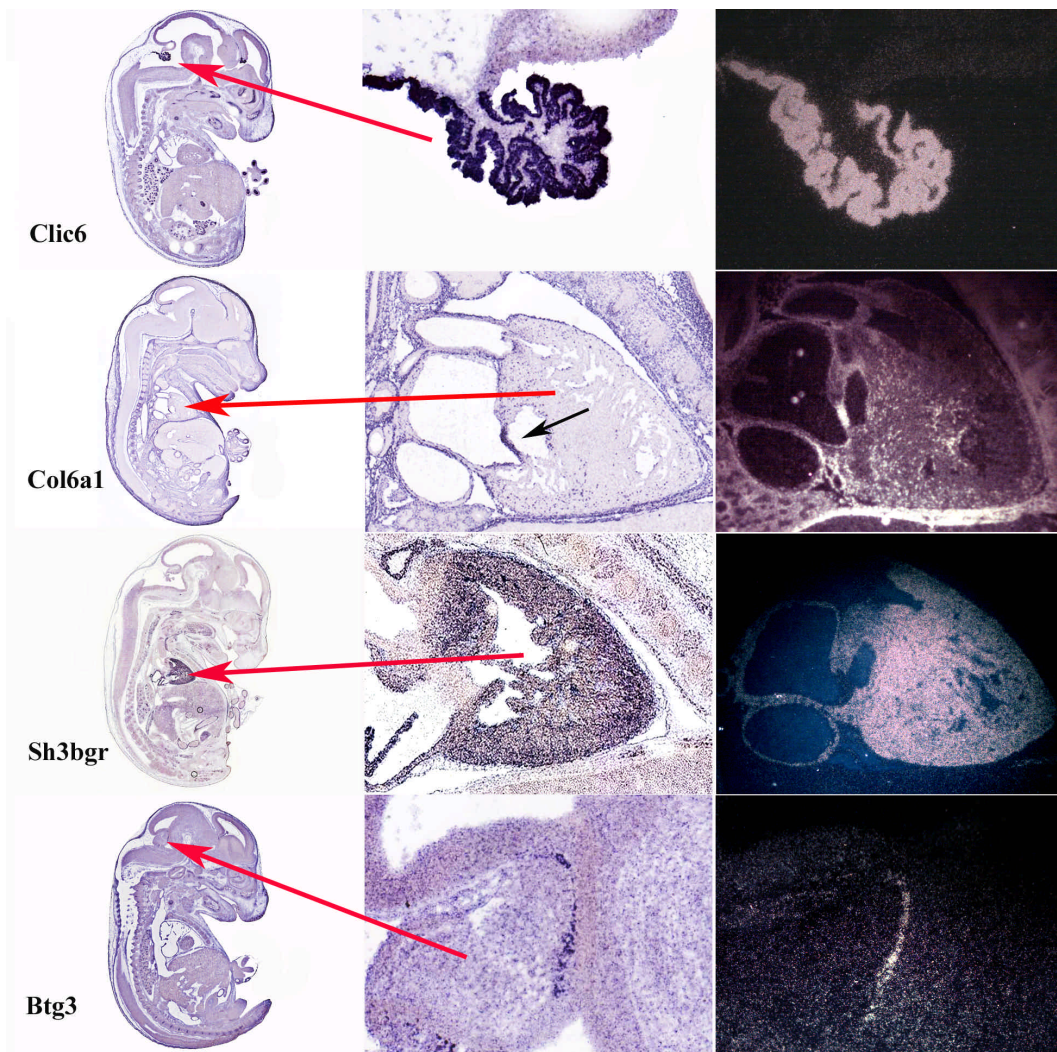


Figure 12. Radioactive vs. non-radioactive ISH panel 1. Starting from the first row from top to bottom; Row 1; *Clic6* is locally expressed in the choroid plexus. Row 2; *Col6a1* shows expression in the mitral valve of the heart (black arrow). Row 3; *Sh3bgr* (*Sh31da*) shows uniform expression in cardiac smooth muscle. Row 4; *Btg3* is in a few cells migrating along the cerebellum. The first column shows the non-radioactive RNA ISH. The second column is a high-power view of a specific region indicated by a red arrow in the left column. The third column shows the results of radioactive RNA ISH for the selected region.

Taken together, our investigations point out that the TSA non-radioactive procedure was certainly equivalent to the radioactive ISH. In fact, the procedure was in several aspects superior. First, it has an advantage where the signal was observed with cellular resolution. This was difficult to achieve with the radioactive procedure because the signal arises through autoradiography in an emulsion that overlies the section. Second, the procedure was much faster. Even in the case of weakly expressed genes, the expression data was viewed in less than 24 hours after starting the analysis. Third, the inherent danger of

working with radioactive compounds and the problems related to their disposal are circumvented.

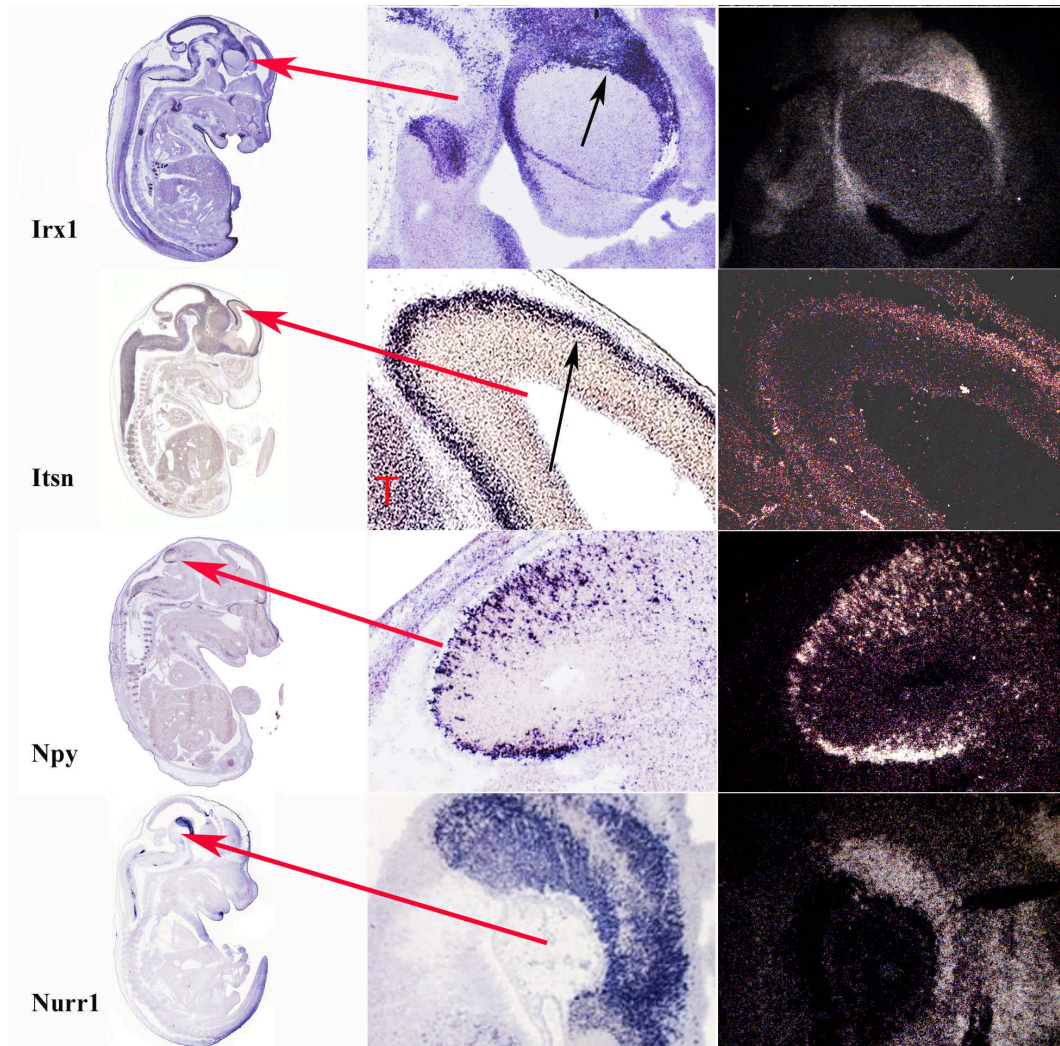


Figure 13. Radioactive vs. non-radioactive ISH panel 2. Row 1; *Irx1* is expressed within the tegmentum (black arrow). Row 2; *Itsn* shows very specific expression in the cortical plate (black arrow) and a relatively weaker signal in the thalamus (T). Row 3; *Npy* is expressed in isolated cells within the *inferior colliculus*. Row 4; *Nurr1* is expressed in the midbrain hindbrain junction, most probably within the dopaminergic neuron precursors. The first column shows the non-radioactive RNA ISH. The second column is a high-power view of a specific region indicated by a red arrow in the left column. The third column shows the results of radioactive RNA ISH for the selected region.

3.1.3 Quantification of the copy number of *Dscr3* transcripts in extracted mRNA from P7 brain sample using real-time PCR

To assess the sensitivity (copy number per cell) of the ISH using hapten labeled RNA, the detection limit of this method was quantified by measuring the number of transcripts present in a defined tissue volume using quantitative PCR. For this purpose a uniformly expressed gene (*Dscr3*) was chosen, that in the ISH analysis gave a weak but detectable

signal (Figure 14A, B). In a first step, three postnatal P7 brains of littermates were dissected and a piece of brain tissue was collected corresponding to a near-rostral quarter portion of the brain (see Figure 14C). One brain was embedded in OCT for later sectioning with a cryostat. The two other specimens were used for DNA and RNA extraction. mRNA was isolated from the brain tissue using an oligo dT procedure. The DNA was isolated using standard extraction techniques. The quantitative PCR will allow an estimate of the amount of *Dscr3* transcripts, while DNA quantification permits an estimate of cell number in the tissue piece. Finally, the microscopic analysis of sections prepared from the remaining tissue piece allowed a second, independent estimate of cell number. mRNA was isolated from brain tissue using the Dynabeads Oligo (dT)25 procedure. The Ambion Competitor Construction Kit was used to generate a *Dscr3* RNA standard with a known number of RNA copies per unit volume. The synthetic RNA referred to as 'competitor RNA' was purified, the concentration measured and the copy number was calculated using the following formula:

$$\frac{\text{purified competitor RNA (CPM/}\mu\text{L)}}{\text{transcription reaction (CPM/}\mu\text{L)}} \times \frac{3 \times 10^{14} \text{ molecules of rATP/}\mu\text{L}}{\text{\# of ATP's in the competitor RNA}} = \text{copies of competitor RNA/}\mu\text{L}$$

$$\frac{5578.1}{339452.2} \times \frac{3 \times 10^{14}}{115} = 4.2869 \times 10^{10} \text{ copies of competitor RNA/}\mu\text{L}$$

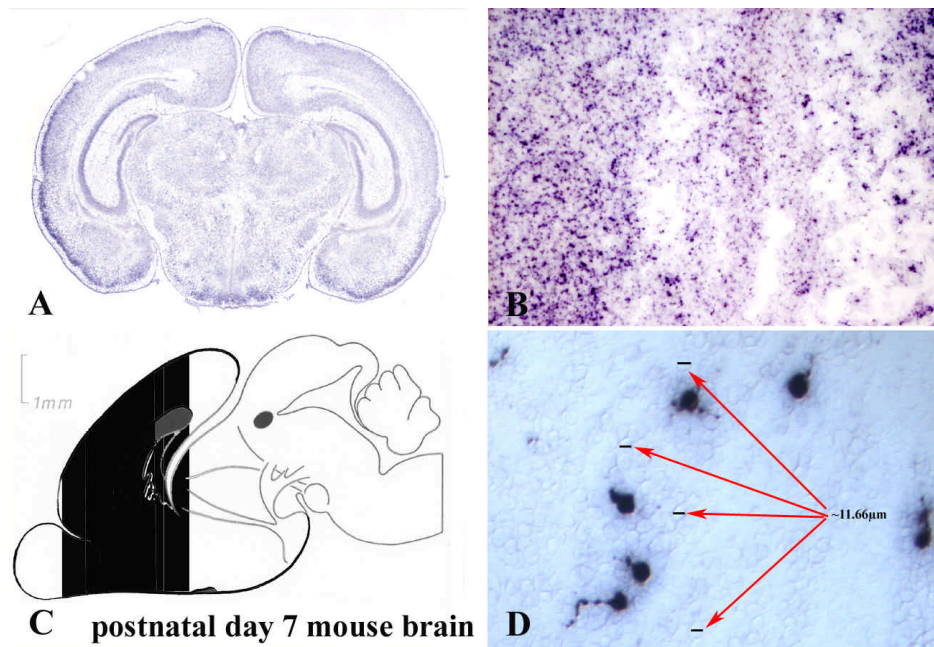


Figure 14. Quantification of a low expressed gene. (A) A coronal section of a P7 mouse brain is represented. This tissue was hybridized with the *Dscr3* RNA probe. (B) Magnification of A. It is easy to see the granular precipitation of NBT-BCIP arising from alkaline phosphatase activity. This signals that there are *Dscr3* transcripts within the vicinity of the granule. (C) A diagram of a

postnatal P7 mouse brain where the near-rostral portion of the brain corresponding to the collected piece of tissue is painted in black. This piece of brain tissue was removed for use in the described experiments. (D) A cell with $\sim 11\mu\text{m}$ diameter is shown. The lines depict the diameter of the cells. To better see the cell, a section subjected to ISH with a *NPY* probe was added.

Real-time PCR amplification was carried out in an iCycler iQ (Biorad) to determine the mRNA copy number for the *Dscr3* gene in the tissue sample. A preliminary analysis using the *Dscr3* competitor RNA as a standard at 10 fold serial dilutions was followed by a second, optimized analysis in which the concentration of the standard encompassed that of the RNA obtained from the brain extract. The preliminary analysis yielded approximately $1.5 \times 10^{+07}$ copies of *Dscr3* RNA per $2\mu\text{L}$ (data not shown). The optimized analysis was carried out using a *Dscr3* RNA standard bracketing this concentration. This real-time PCR showed that the unknown sample contained $1.4 \pm 0.5 \times 10^{+07}$ copies per $2\mu\text{L}$ (Figure 15). Since the total volume of the mRNA extract was $30\mu\text{l}$ and the extraction efficiency was approximately 50% (as determined by OD measurement), the total number of *Dscr3* mRNAs in the piece of brain used was estimated to be $4.2 \times 10^{+08}$ copies.

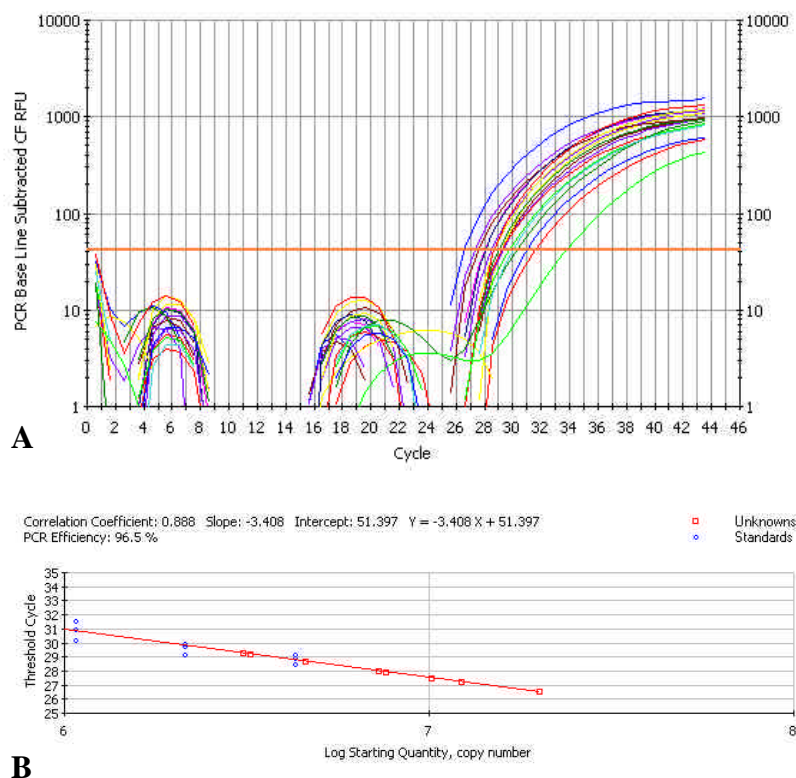


Figure 15. Optimized quantitative PCR analysis to determine the copy number of transcript from the *Dscr3* gene. (A) The chart showing a plot for each PCR reaction based upon its relative fluorescence unit (RFU). Baseline subtraction is made at the lowest point where linear amplification of the PCR reactions is occurring. (B) The standard curve with *Dscr3* competitor RNA was obtained by plotting of signals in (A) for three concentrations of standards (blue points, each concentration was examined in three replicates). Similarly, the unknowns was analyzed in serial dilutions and parallel replicates. Using the Icyler IQ Optical System statistical analysis

software (version 3.0a) the quantitative PCR yielded approximately $1.4 \times 10^{+07}$ copies of *Dscr3* RNA per 2 μ L of unknown sample.

To determine the number of *Dscr3* mRNA molecules per cell, we next determined the cell number in the tissue samples. This was done in two different ways. First, we estimated the number of cells using the DNA content per cell. Second, we estimated the cell number per section in a 20 μ m thick section and extrapolated this estimate to the number of cells in the entire piece of brain tissue that was used for mRNA extraction.

DNA-based cell count: The DNA content of the brain piece was determined as 333.36 μ g (Figure 14C). Since the diploid DNA content of a mouse cell is 3pg, the brain piece contained $\sim 1.1 \times 10^8$ cells.

Morphometry based cell count: First, the tissue piece was sectioned with a cryostat at a thickness of 20 μ m. Then the area of each section was determined in the microscope. On the average, each cell measures approximately $\sim 1250 \mu\text{m}^3$ [$(\sim 11.6)^3 \mu\text{m}^3$] (Figure 14D). The tissue piece consisted of approximately 256 sections, and the tissue volume was $\sim 1.7 \times 10^{11} \mu\text{m}^3$. Therefore, these measurements yield $\sim 1.36 \times 10^8$ cells in the tissue piece.

The cell count estimates ranging between 1.1 and 1.36×10^8 (average is 1.23×10^8) and the copy number of *Dscr3* mRNA molecules determined by quantitative PCR now allow one to estimate the number of mRNA molecules present per cell. Recall that the tissue sample contained $4.2 \times 10^{+08}$ copies of *Dscr3* mRNA. Thus there are

$$\frac{4.2 \times 10^8 \text{ copies}}{1.23 \times 10^8 \text{ cells}} = \sim 3.4 \text{ copies per cell}$$

We conclude that a weak signal as exemplified by *Dscr3* (Figure 14 A, B) was caused by as few as 3 molecules of *Dscr3* mRNA. Therefore, the non-radioactive ISH procedure used for the present work is very sensitive similar to the sensitivity reported for ^{35}S based ISH (Speel et al. 1998).

3.1.4 Improved ISH performance using optimized probes

In some cases, instead of using commercially available clones as templates (such as was done for most probes used in this study), hapten labeled RNA probes were generated by RT-PCR based procedure (described in Section 2.2.13). The improvement in the results was outstanding for the HC21 orthologue mouse genes *IFNGR2*, *RUNX1*, *DYRK1A*,

GRIK1, and *JAM2* when probes were used that were derived from a specifically designed PCR template. The expression patterns revealed by the old and new (RT-PCR based) mouse *Ifngr2* gene probe template are shown in Figure 16. These data show the importance of proper probe design. The new probe does not contain a poly A tail, is 200 nucleotides longer and contains less 3 prime untranslated region compared to the old probe.

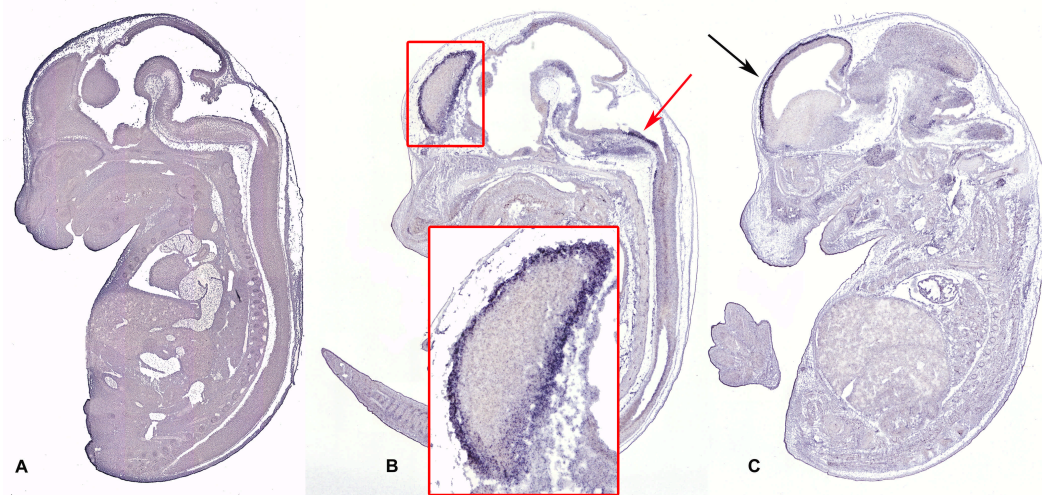


Figure 16. *Ifngr2* gene expression detected by two different probes. (A) The specimen hybridized with the old probe (made from clone template) had high ubiquitous expression with no expression (B-C) The specimen analyzed with the RNA probe made from the specifically designed template showed regional expression in the cortical plate (red box and black arrow) and in the medulla (red arrow).

3.2 Gene expression atlas of the mouse orthologues of the majority of human chromosome 21 genes

In the previous section the enhancement that increases the sensitivity of the ISH protocol we use was illustrated. We also provided an estimate of the number of transcripts this method can detect. Finally the non-radioactive protocol was compared to a conventional radioactive ISH protocol. After assuring that the protocols were optimized to a degree where the ISH protocol is now useful, the method was tested by cataloguing the expression patterns of all the mouse orthologues of the HC21 genes. This effort thus represents the first large-scale application of the automated non-radioactive ISH technology to a biological problem.

199 cDNA fragments representing 150 mouse orthologues of the human chromosome 21 genes were isolated for ISH experiments that were carried out on sagittal sections of embryonic day 14.5 mouse embryos. Due to the fact that nearly all the organs and tissues

have completed their development makes this a very important stage when analyzing genes expressed in DS relevant organs (brain, heart, skeleton, gut, thymus, pancreas). Therefore this high resolution expression “atlas” of an entire human chromosome is an important step towards the understanding of gene function and of the pathogenetic mechanisms in DS.

The fact that there was a vast amount of data produced in the course of the expression analysis of the genes of a whole chromosome, it was important to be able to present the data in a meaningful manner. Aiming to provide an interface that would satisfy the needs of users with different backgrounds and expertises, an interface capable of achieving this goal was designed. Therefore as a first step it was necessary to verbally annotate the expression patterns of the mouse orthologues of the HC21 genes. This way scientists working to find a solution for congenital heart defects would have the opportunity to easily overview genes expressed within the heart at the investigated stage. The annotation process was detailed as far as possible by dividing the organs into substructures. This was done by trying to reach a balance between a detailed and informative annotation and time constraints inherent to the annotation effort. Highest priority was given to trying to include the most relevant and important DS organs. Then the annotation of the genes was presented in a table, that preserves the chromosomal order of the HC21 genes. A web page was designed and made public for the easy access of expression data (images, annotation tables and experimental metadata) for the “human chromosome 21 gene expression atlas”. The web site “www.tigem.it/ch21exp” is described in appendix 7.1. Table 10 represents an example annotation table that can be found on this website. This table provides information on where the *Btg3* gene is expressed in the E14.5 mouse embryo. The annotation criteria can be explained as follows: if the signal was regional, like *Itsn* expression in the cortical plate of the telencephalon it was assigned an “r” for regional expression (Figure 17A). When the gene was expressed everywhere in the relevant tissue or structure in question it was considered ubiquitous “u” (Figure 17B). In the case for the *H2bfs* gene, where isolated “positive” cells across the tissue were observed, the gene was assigned an “i” for isolated expression (Figure 17C). The signal strength was scored according to the following levels of expression: three stars (***) for strong expression as was observed for the *Itsn* gene in the cortical plate (Figure 17D, black arrow), two stars for medium expression (which was lower compared to the strong expression in the cortical plate of the same section) seen for *Itsn* expression in the thalamus (Figure 17D, red box). One star for weak expression, such as that observed for the *Dscr3* gene in the pons (Figure 17E). When the gene (*Cxadr* expression in medulla)

did not show expression it was not scored (Figure 17F). The *Btg3* gene has also weak ubiquitous expression along with patterning. This is why it has an “ubiquitous **” sign in the overall tissue expression box. The whole table was also shaded in light grey which is darker when the signal strength is stronger (e.g. ubiquitous **).

Table 10. Example annotation of the *Btg3* gene.

gene name	Btg3
ISH number	06/02-1
Specimen	Ex75.1
Tissue quality	**
Overall Tissue expression	ubiquitous *
Forebrain	** R
Midbrain	
Hindbrain	** R
Cerebellum	*** I
Spinal cord	*** I
Cranial ganglia	
Dorsal root ganglia	
Parasympathetic ganglia	
Eye	
Nasal/olfactory epithelium	
Cochlea	
Skeleton (cartilage)	
Lungs	
Atrium	
Ventricles	
Arterial system	
Gut	
Bladder	
Kidney	
Testis	
Ovary	
Salivary	
Spleen	
Pancreas	
Adrenal	
Thyroid	
Thymus	
Liver	
Skin	
Whisker follicles	
Muscles	
Mesenchyme	
Tongue	
Limbs	
Tail	

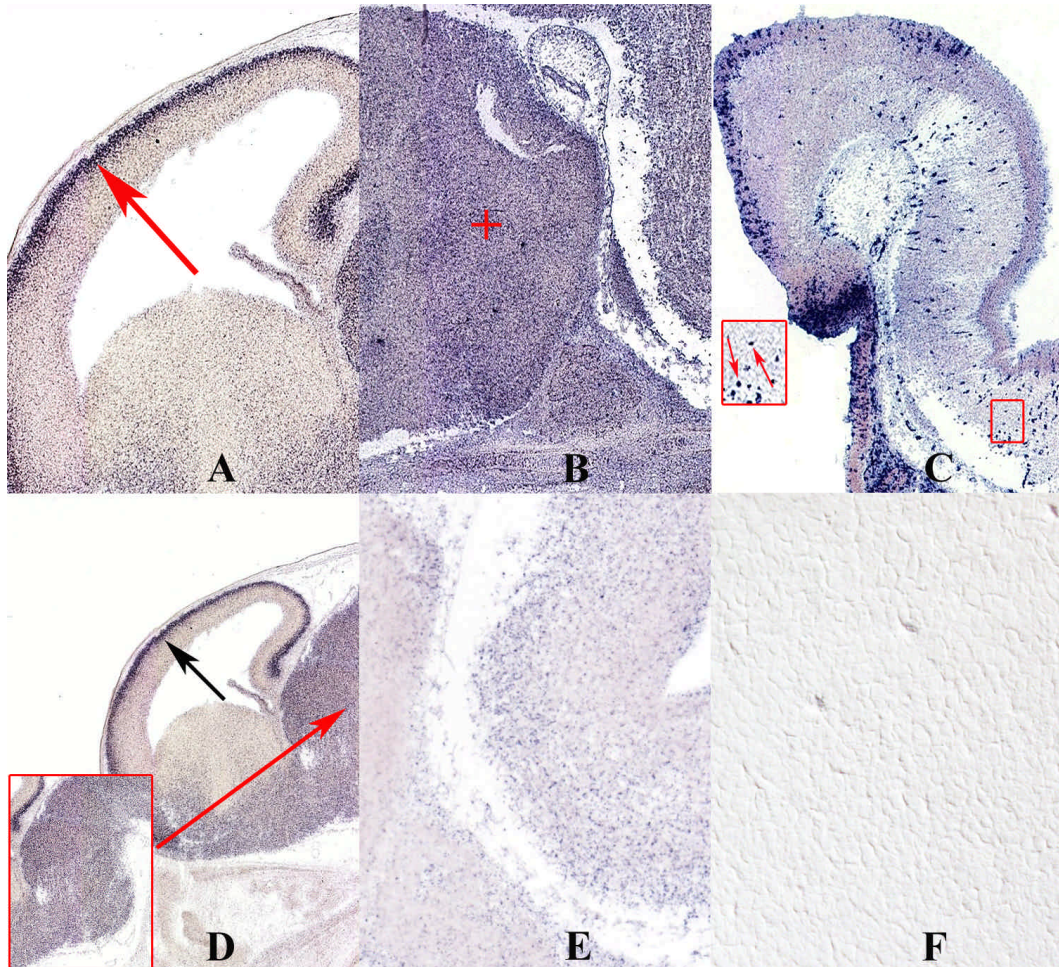


Figure 17. Annotation of the signal intensity and pattern observed for expression. (A) *Itsn* (*intersectin*) showing strong (***) regional (R) expression in the cortical plate (red arrow) of the telencephalon. (B) *Nnp1* expression in the hypothalamus exemplifies strong and ubiquitous “u” expression. (C) The *H2bfs* gene is expressed in isolated cells across the tissue and was assigned an “i” for isolated expression (red arrows pointing to isolated cells in the red box). (D) The signal strength was scored three stars (***) for strong expression, as was observed for the *Itsn* gene in the cortical plate (black arrow), two stars for medium expression for *Itsn* expression in the thalamus (red box). (E) Two stars for medium expression, such as that observed for the *Itsn* gene in the thalamus. (F) One star for weak expression, such as that observed for the *Dscr3* gene in the pons. When the gene (*Cxadr* expression in medulla) did not show expression it was not scored.

3.2.1 Statistics

The analysis of the 6500 tissue sections generated for this HC21 atlas provides the following statistical overview. Expression was detected for 64% of the genes examined at embryonic day 14.5. Regional gene expression was observed in 42% of the genes (see Figure 18). The highest number of genes with a restricted expression pattern were observed in the brain and the gut. Ubiquitous expression was observed in 13% of the cases. Genes with both weak ubiquitous and strong regional expression were observed in 9% of the cases. No expression was detected for 36% of genes.

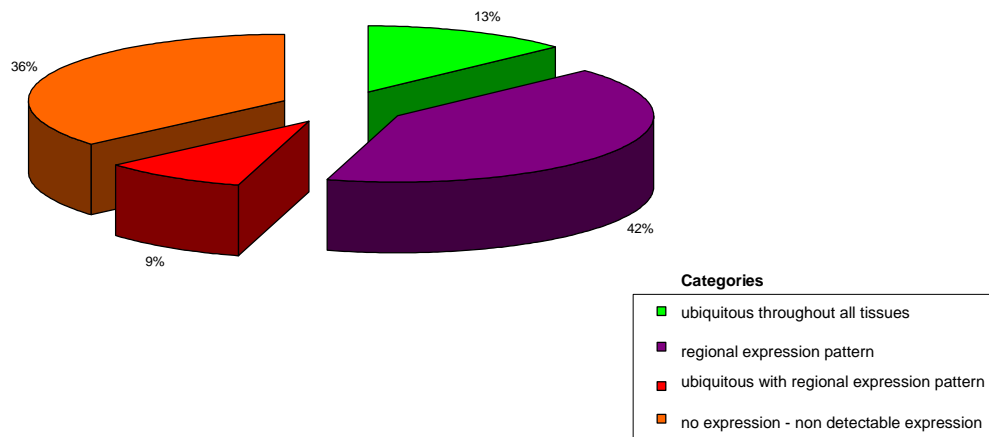


Figure 18. Distribution of expression patterns at embryonic day 14.5. In this pie-chart each slice represents the percentage of genes belonging to the four categories of expression patterns observed.

The resolution of the technology at hand, allows the study of genes showing regional expression within organs and tissues. The well developed organ structure of the embryonic day 14.5 embryo made it possible to visualize the expression of the human chromosome 21 mouse orthologue genes within DS affected organs. Brain, heart, gut, limbs, thymus, pancreas are organs that have been related to the DS phenotype. Figure 19 is a pie chart showing genes that have regional expression within these organs. The highest percentage of regional expression was observed in the brain (44%) for the human chromosome 21 orthologue mouse genes. This was followed by the gut with 17% of the genes, the heart, limbs and the thymus have an equal slice of 10% and the pancreas had the lowest slice with 9% of the genes. In the next section expression in these tissues will be discussed in relation to the DS pathology.

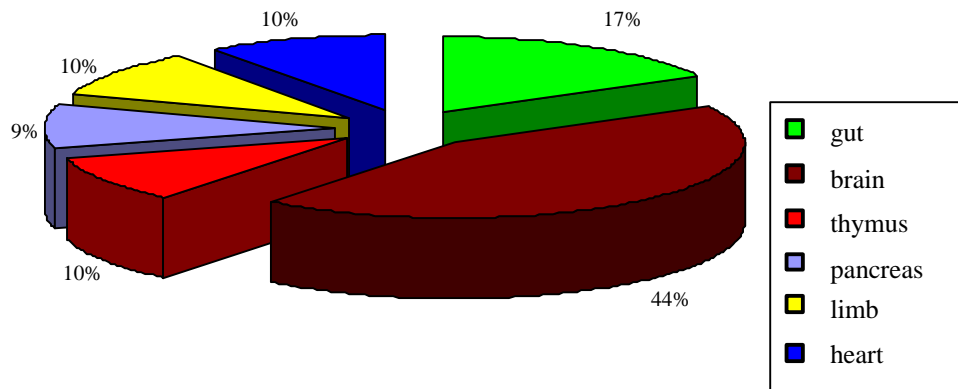


Figure 19. Regional expression observed in Down syndrome affected organs. The slices of the pie-chart correspond to the percentage of genes showing regional expression in DS related organs at embryonic day 14.5 via RNA ISH.

3.2.2 Genes expressed within the brain

Certain DS phenotypes are not observed in every single DS patient. For example, only 40% of DS patients have congenital heart defects. By contrast, mental retardation is a phenotype which is present in all DS individuals. This phenotype is associated with the brain. The brain is involved in carrying out mental functions, where it is known that Down patients have a lower intelligence quotient (IQ) compared to normal individuals. Mental retardation, is the below normal level of intellectual functioning and the decreased capability for adaptive behaviour. Stereological counting techniques have revealed that the second phase of cortical development and the emergence of cortical lamination are both delayed and disorganized in DS fetuses. Striking cellular changes in the DS brain are observed in newborns and infants. They have shorter dendrites, a decreased number of spines with altered morphology and a defective cortical layering. Studies have also shown delayed myelination, fewer neurons, lower neuronal density, and delayed synaptogenesis (Engidawork and Lubec 2003).

Another defect involving the brain of DS patients is that they show the pathological changes observed in people with Alzheimer disease. Although Alzheimer is usually observed at old age in normal individuals, DS patients show the disease symptoms after their third decade. Early onset Alzheimer disease is a very important disease related to neuronal degeneration occurring through the accumulation of amyloid plaques. The above mentioned factors, which have been associated to defects arising in the brain, encouraged the necessity of a panel that represents genes expressed within the developing brain at embryonic day 14.5 (Figures 20-26).

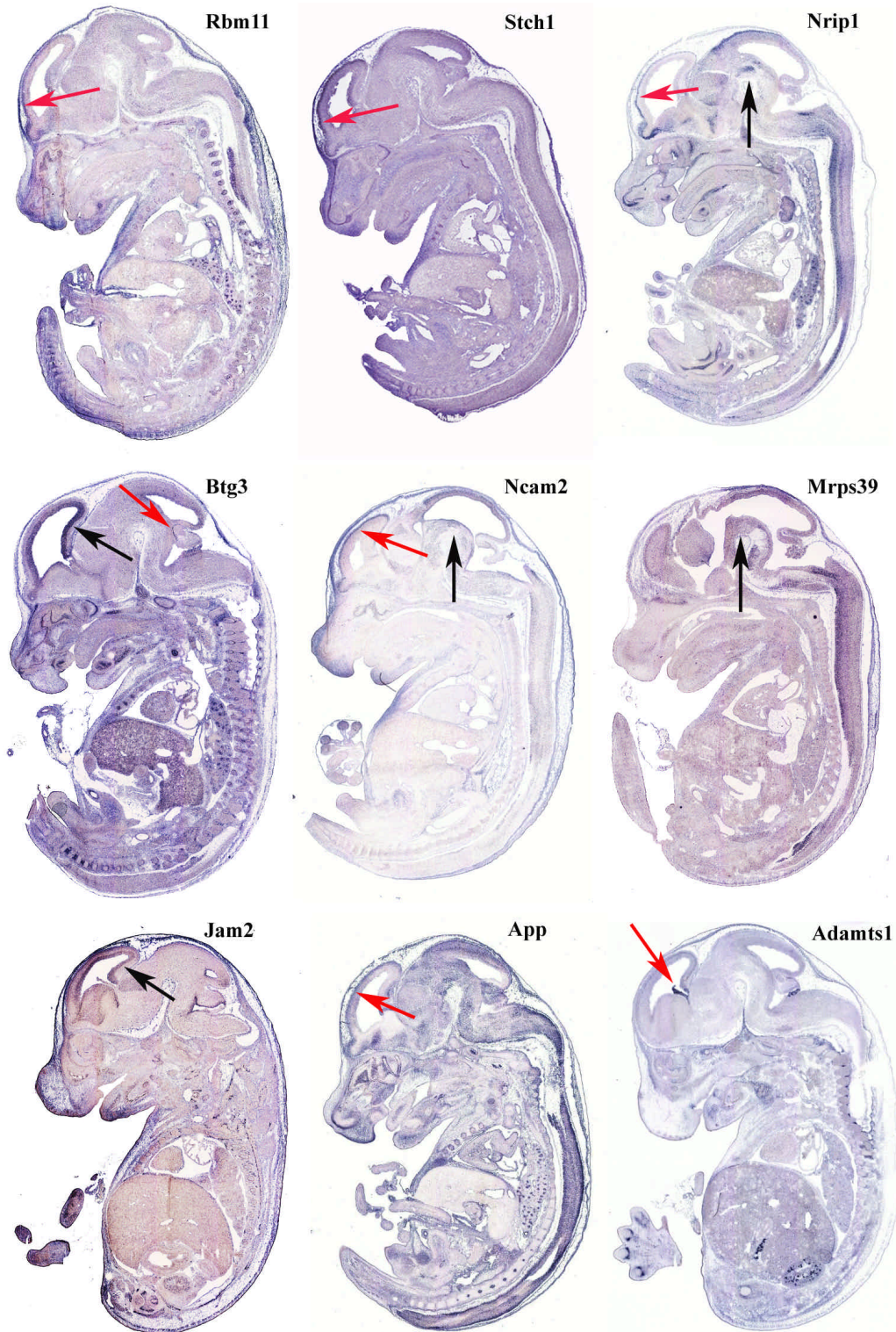


Figure 20. First panel of genes expressed within the brain. Most of the genes in the panel were expressed in the cortical plate *Rbm11* (red arrow), *Stch1* (red arrow), *Nrip1* (red arrow), *Ncam2* (red arrow), *App* (red arrow) and the ventricular zone *Btg3* (black arrow), *Jam2* (black arrow) of the telencephalon. *Btg3* also showed expression in a few specific cells migrating down to the cerebellum from the midbrain (red arrow). *Adamts1* was locally expressed within the choroid plexus (red arrow). *Ncam2* (black arrow), *Mrps39* (black arrow) and *Nrip1* (black arrow) also showed expression in the midbrain hindbrain junction.

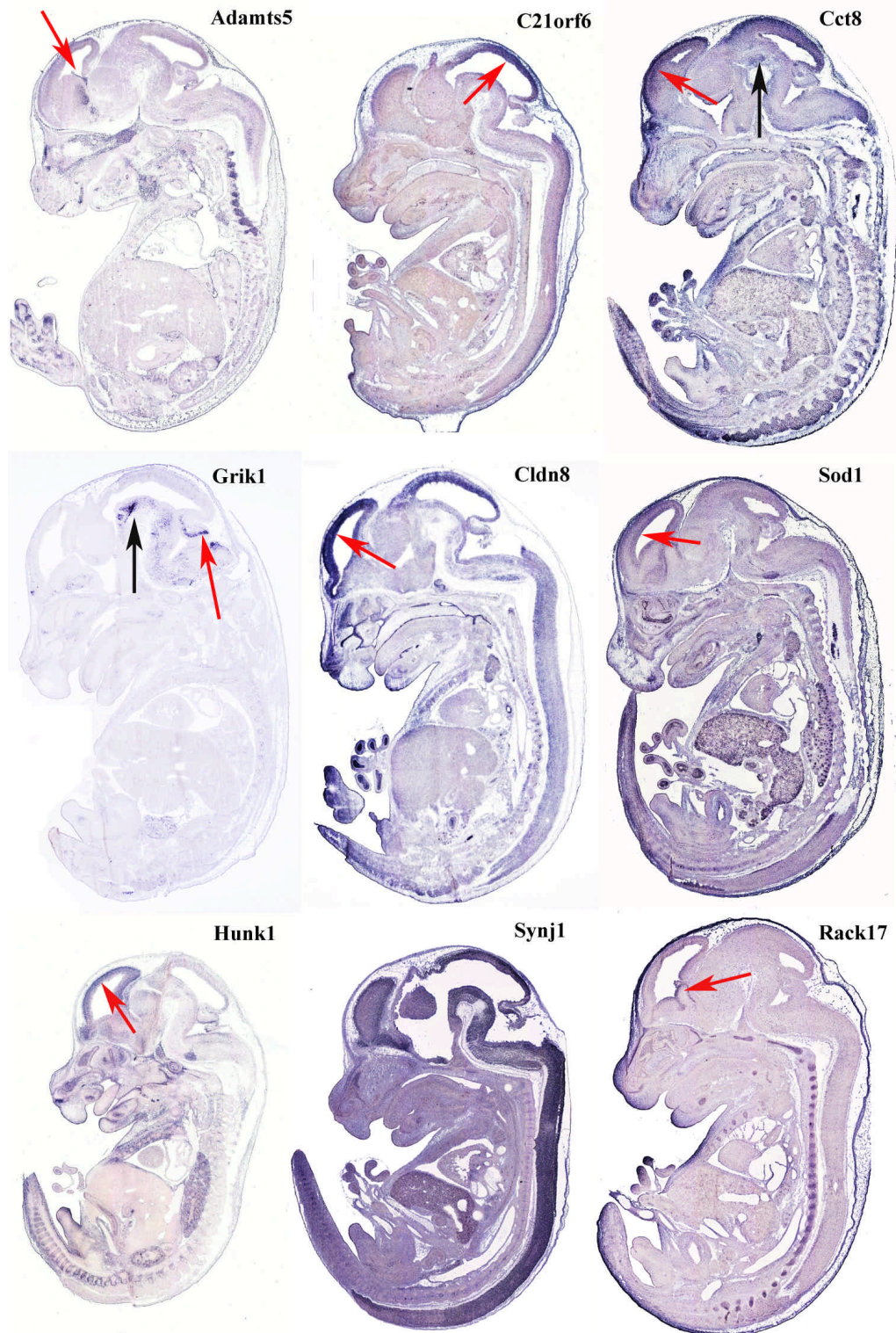


Figure 21. Second panel of genes expressed within the brain. A number of the genes in the panel were expressed in the cortical plate (*Cct8* (red arrow), *Cldn8* (red arrow), *Hunk1* (red arrow)) or the ventricular zone (*Sod1* (red arrow) and *Rack17* (red arrow)) of the telencephalon. *Adamts5* (red arrow) was locally expressed within the choroid plexus. *Grik1* shows local expression in the developing cerebellum (red arrow) and in the midbrain hindbrain junction (black arrow). *Cct8* (black arrow) also shows expression in the midbrain hindbrain junction. *Synj1* shows broad expression with patterning where transcript levels are very high in the brain.

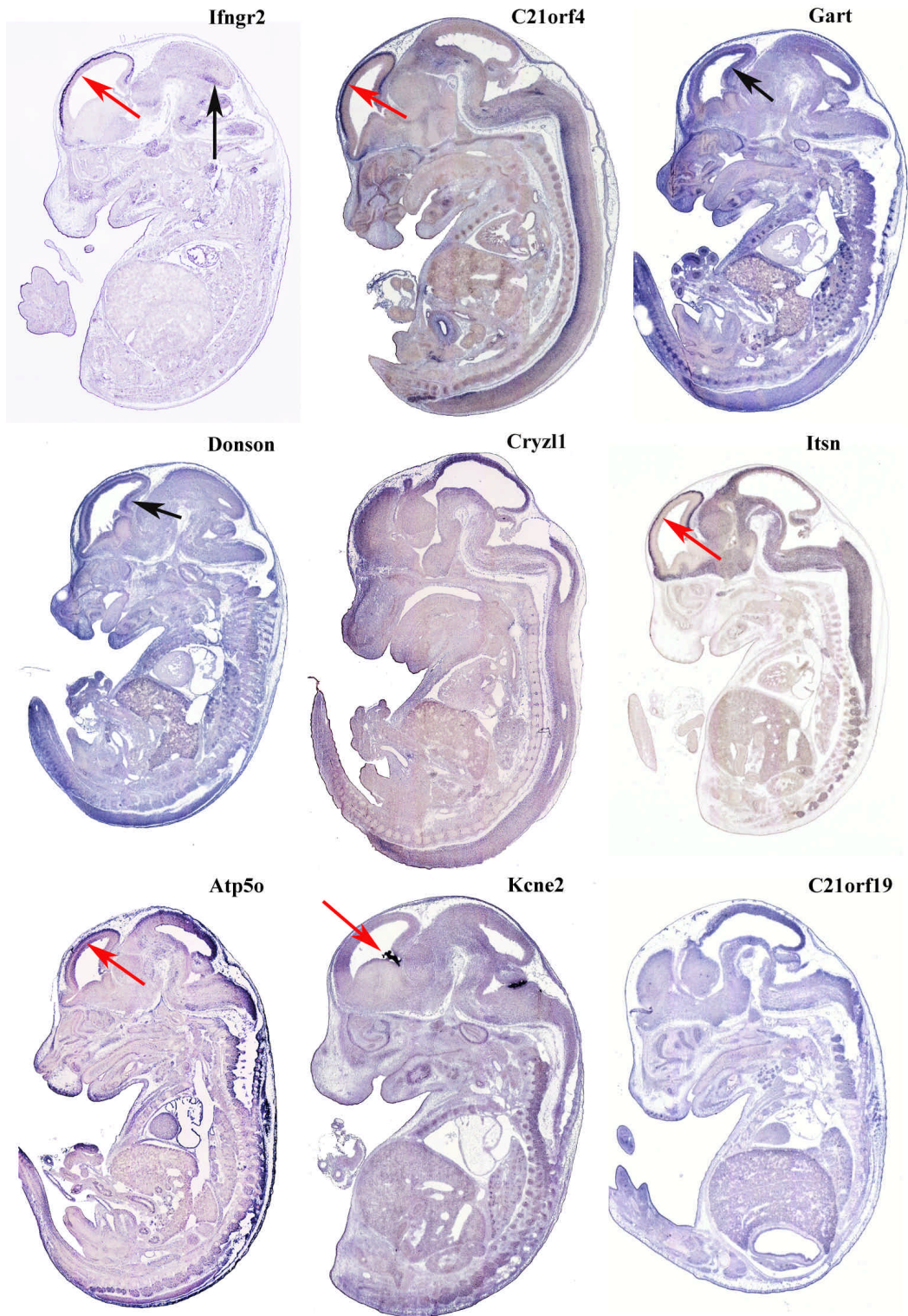


Figure 22. Third panel of genes expressed within the brain. The genes *Ifngr2* (red arrow), *C21orf4* (red arrow), *Itsn* (red arrow), *Atp5o* (red arrow) which are represented in the figure were expressed in the cortical plate. The genes *Gart* and *Donson* were expressed in the ventricular zone of the telencephalon. *Kcne2* was expressed within the choroid plexus (red arrow). *Cryz1* shows broad expression in the whole brain. *Ifngr2* was also expressed in a streak of cells in the midbrain (black arrow).

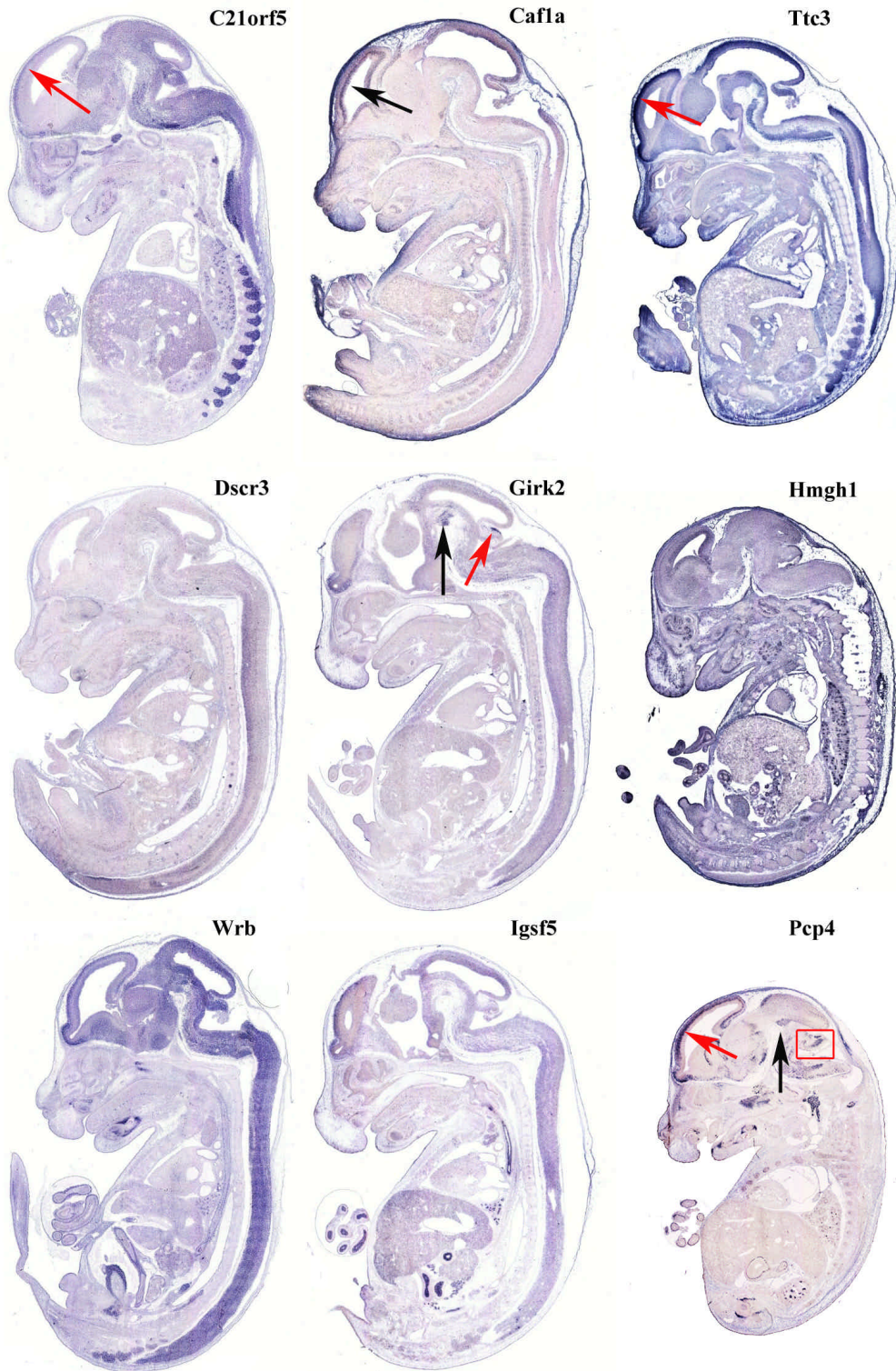


Figure 23 Fourth panel of genes expressed within the brain. *Hmgh1*, *Dscr3*, *Igsf5* and *Wrb* show uniform expression in the brain. *Pcp4* shows expression in the cortical plate (red arrow), developing cerebellum (red box) and in the midbrain hindbrain junction (black arrow). *Girk2* was also expressed in the developing cerebellum (red arrow) and the midbrain hindbrain junction (black arrow). *C21orf5* (red arrow) and *Ttc3* (red arrow) were expressed in the cortical plate. *Caf1a* (black arrow) was expressed in the ventricular zone of the telencephalon.

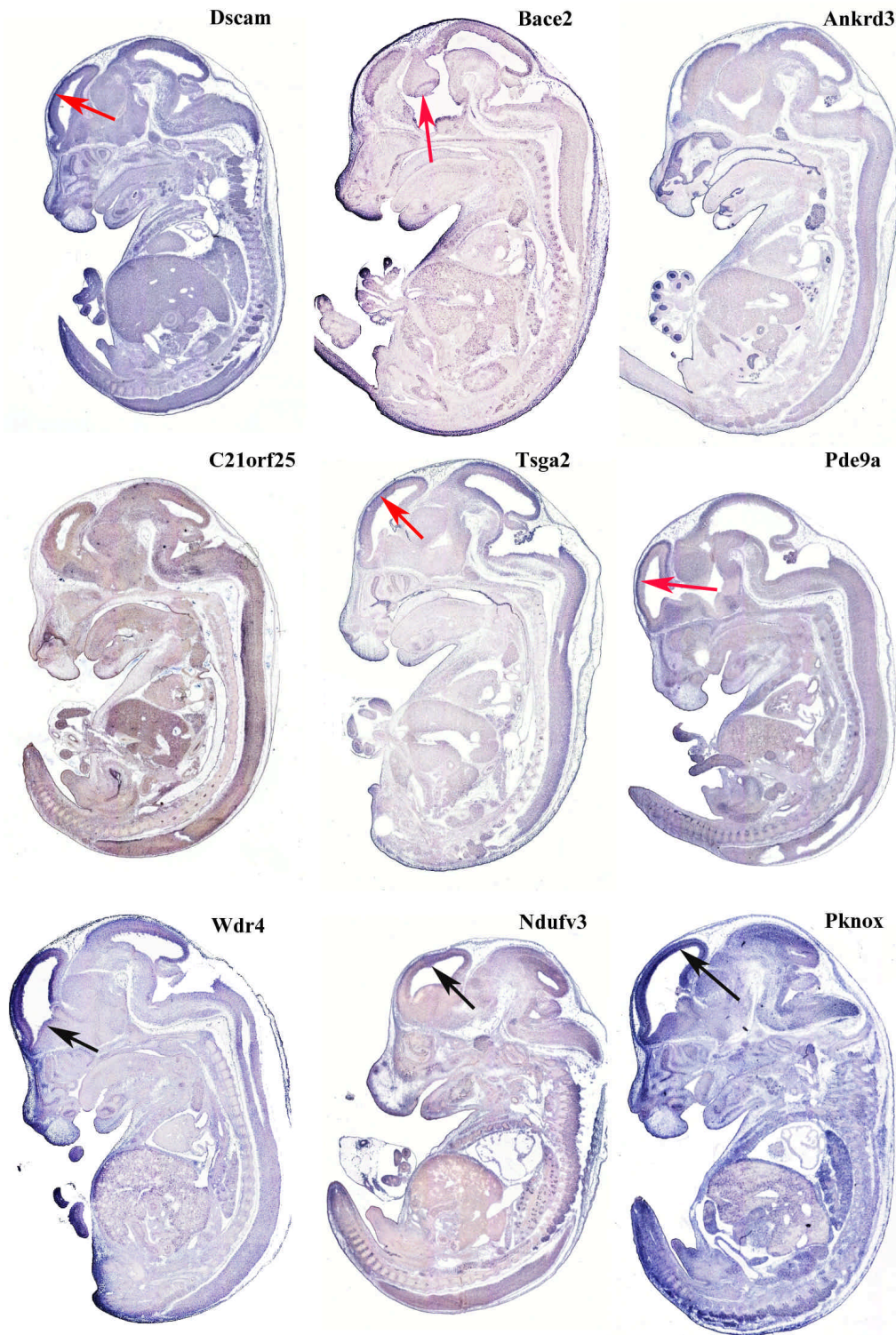


Figure 24. Fifth panel of genes expressed within the brain. *Dscam* (red arrow), *Pde9a* (red arrow) and *Tsga2* (red arrow) show expression in the cortical plate. *Pknox* (black arrow), *Ndufv3* (black arrow), *Wdr4* (black arrow) are expressed in the ventricular zone of the telencephalon. *Tsga2* was also expressed within the choroid plexus. *Bace2* is expressed in the thalamus (red arrow). *Ankrd3* show broad expression in the brain.

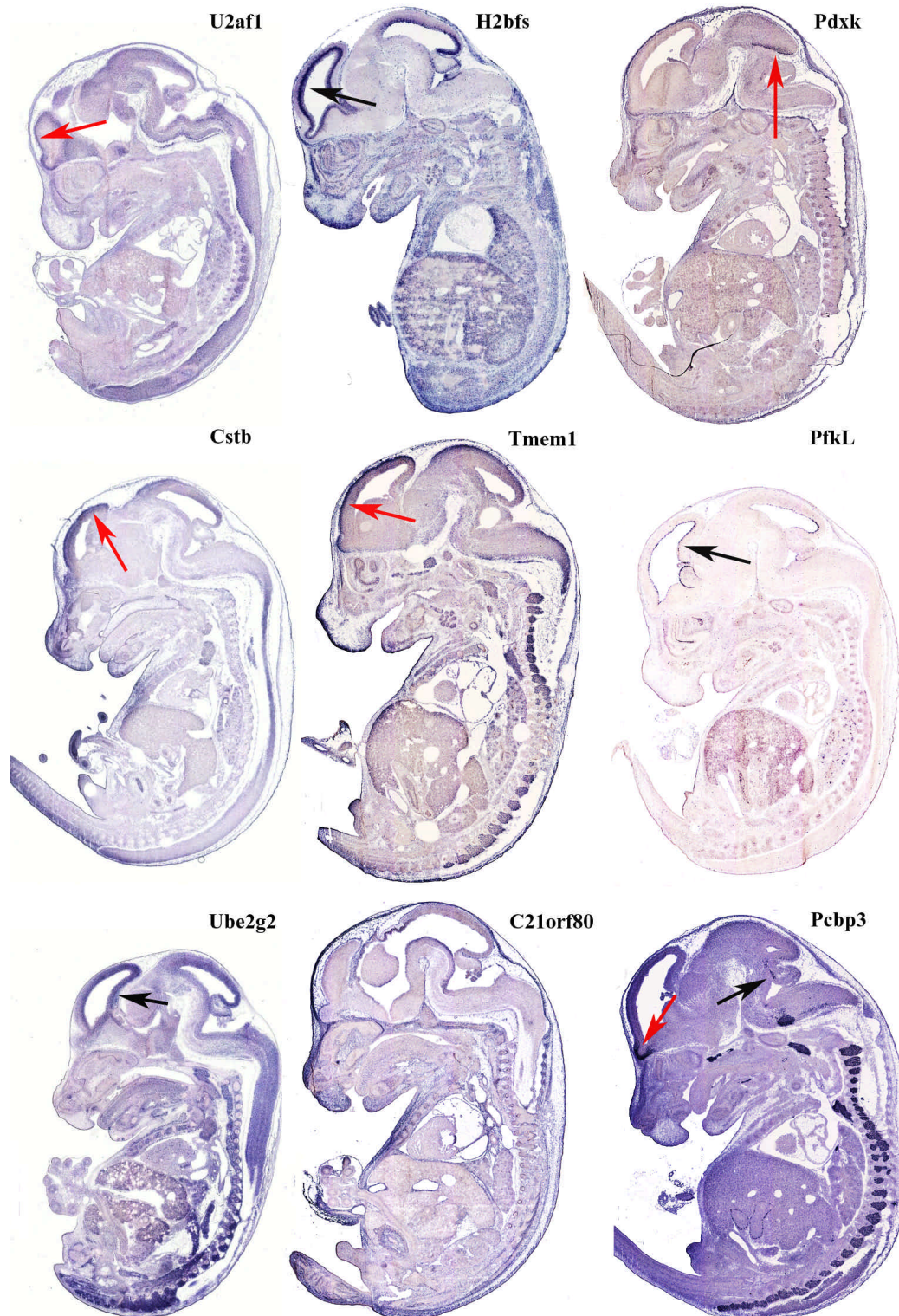


Figure 25. Sixth panel of genes expressed within the brain. *U2af1* (red arrow), *Tmem1* (red arrow) and *Cstb* (red arrow) show expression in the cortical plate. *Ube2g2* (black arrow), *PfkL* (black arrow) and *H2bfs* (black arrow) are expressed within the ventricular zone of the telencephalon. *Pdxk* was expressed in a streak of cells in the midbrain (red arrow). *Pcbp3* was expressed in the dorsal root ganglia, olfactory bulb (red arrow), medulla and a streak of cells stretching into the cerebellum from the midbrain (black arrow).

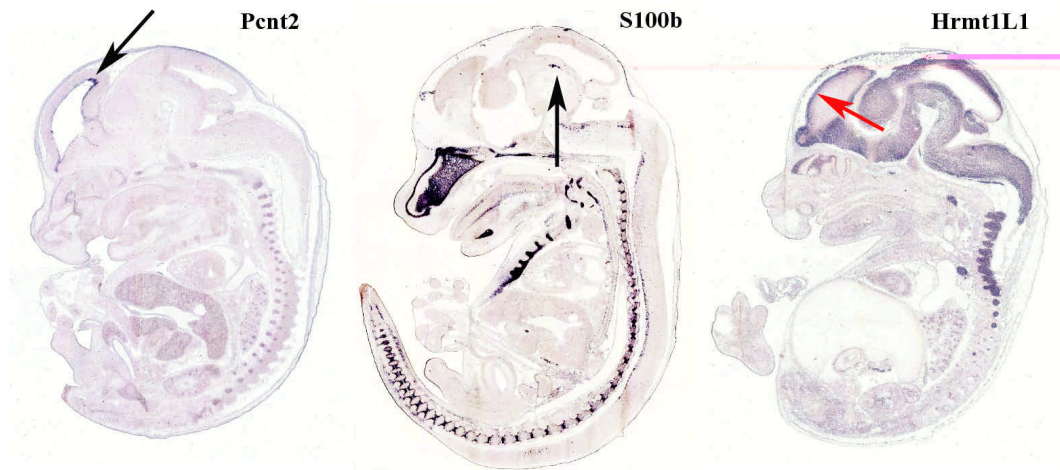


Figure 26. Seventh panel of genes expressed within the brain. *S100b* was expressed in the midbrain hindbrain junction (black arrow) and in the medulla. *Pcnt2* (black arrow) was expressed within the ventricular zone of the telencephalon, while *Hrmt1L1* (red arrow) was expressed within the cortical plate.

3.2.3 Genes expressed within the heart

One of the major causes for death among people with DS is congenital heart disease. Congenital heart disease is observed in 40% of people with DS. 39% of the children affected by DS have been reported to have atrioventricular defects among other cardiac lesions. Apart from congenital heart disease related to DS, chromosome 21 has also been associated to other heart diseases like Long QT syndrome, Ullrich disease, Knobloch syndrome and Bethlem myopathy, and thus provides an important model to link individual genes to pathways controlling heart development. Figure 27 represents the genes showing expression in the heart at embryonic day 14.5. *Pfkl* was strongly expressed in the ventricular wall and atrium in the developing heart. *Kcnj15* and *Adarb1* were expressed in the aortic valve and trunk, while *Kcnj15* was also expressed along the outflow track of the heart and in the superior vena cava. *C21orf7* shows expression within the aortic trunk. *Atp50* and *Sh3bgr* transcripts were detected throughout the heart. *Col18a1* was only detected in cardiac vessels. *Col6a1* and *Col6a2* were strongly expressed in the mitral valve and along the pericardium.

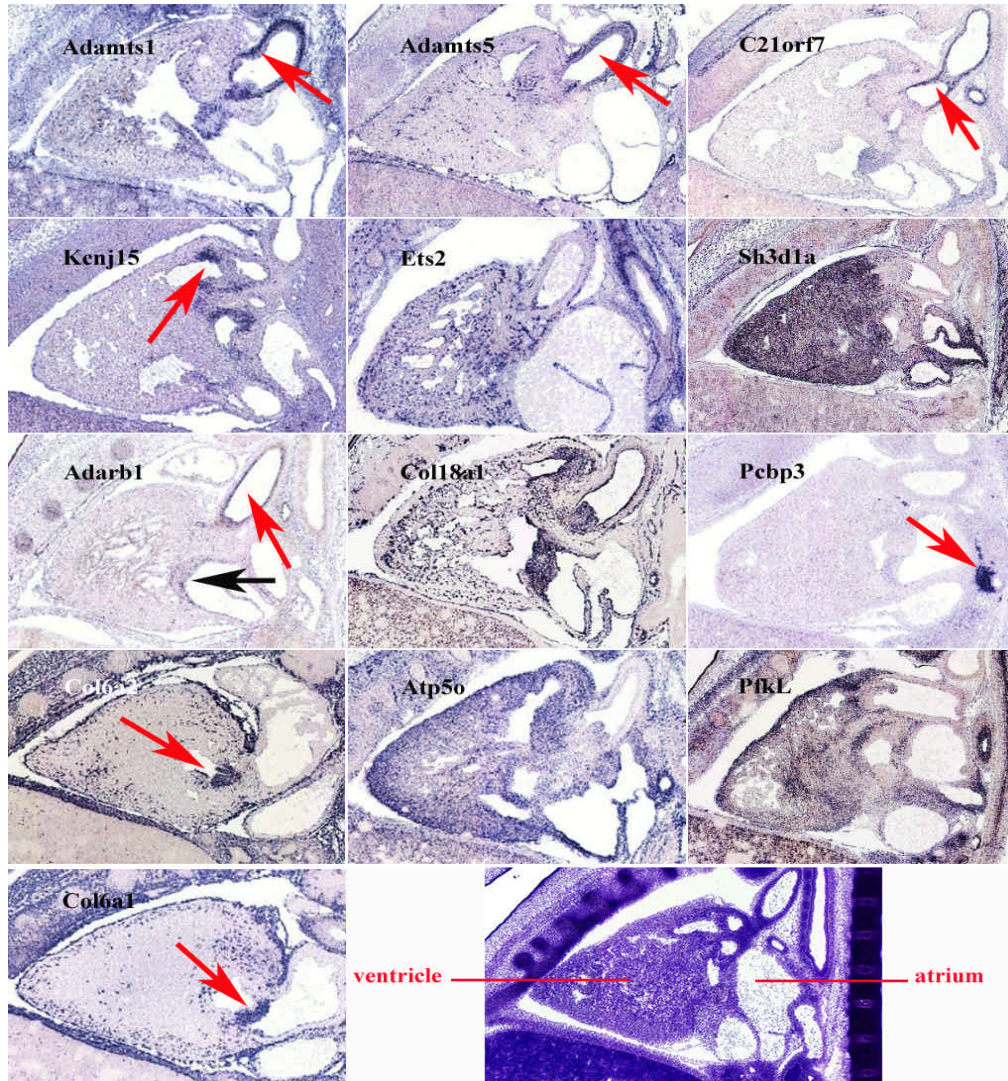


Figure 27. Expression analysis in the developing heart. E14.5 sections were probed for *Adarb1*, *Atp50*, *Col6a1*, *Col18a1*, *Kcnj15*, *PfkL*, and *Sh3d1a* genes. The region portrayed from the embryo is schematically represented in the bottom right panel. *Sh3d1a* (*Sh3bgr*) and *Atp50* transcripts are detected throughout the heart. *Adarb1* mRNA is restricted to the mitral valve (black arrow), aortic valve and the endothelium of the aortic trunk (red arrow). *PfkL* was expressed throughout the heart. *Kcnj15* transcripts are found in the aortic (red arrow) and mitral valves. *Col6a1* was strongly expressed in the mitral valve (red arrow) and along the pericardium, whereas *Col18a1* and *Ets2* expression was seen in the vessels of the heart. *Adamts1* (red arrow), *Adamts5* (red arrow) and *C21orf7* (red arrow) all show expression within the aortic trunk, while *Pcbp3* was expressed within the peripheral nervous system (red arrow).

3.2.4 Genes expressed within the limbs

The most apparent if not the most serious of the DS phenotypes is that all patients have dysmorphic features. DS fetuses exhibit a reduced growth rate of limb long bones during the third trimester of pregnancy. Limb defects like a broad gap between the second and first toe, short broad hands and a non-ossified or hypoplastic middle phalanx of the fifth digit are present in these fetuses (Epstein 1995). For this reason Figure 28 represents genes expressed within the limbs at embryonic day 14.5. *Erg* has strong expression in the

posterior proximal mesoderm and in the joints. *Ets2* is also expressed in the skeletal system. *Adamts1* is expressed in the perichondrium of the developing bones.

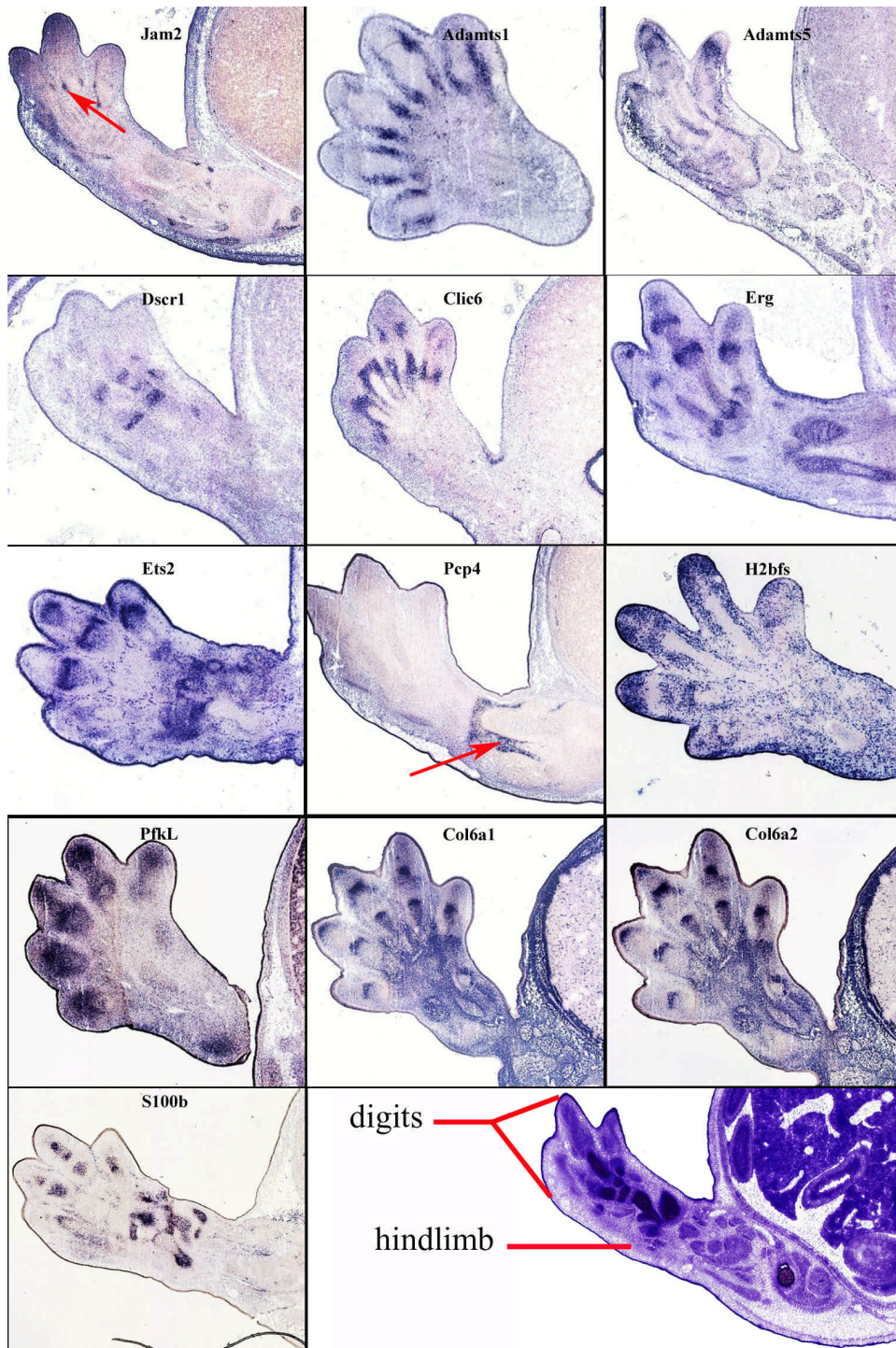


Figure 28. A panel of genes expressed within the developing limb. *Erg* transcripts were present in the posterior proximal mesoderm and in the joints. *Adamts-1*, *Adamts5* and *Clic6* show expression in the perichondrium of the developing digits. *H2bfs* showed expression within the ridges of the digits. *Col6a1*, *Col6a2*, *Ets2* and *PfkL* show specific expression within the joints. While *Jam2* (red arrow) and *Pcp4* (red arrow) were expressed in the tendons of the limbs. *S100b* was expressed within the cartilage of the limbs.

3.2.5 Genes expressed within the digestive tract

The digestive tract is well differentiated at E14.5 and the expression of genes can be easily detected. DS patients have a high frequency for gastrointestinal abnormalities such as duodenal stenosis, Hirschsprung's disease, gastroesophageal reflux and imperforate anus (Epstein 1995). The mouse genes *Atp50*, *Cldn8*, *Clic6*, and *Ets2* were expressed within the endothelium of the duodenum. The *Tff3* and *Sod1* transcripts were present in a subgroup of cells within the gastro duodenal junction region of the pyloric sphincter with some expression in the endothelium of the stomach and the intraperitoneal portion of the midgut, respectively.

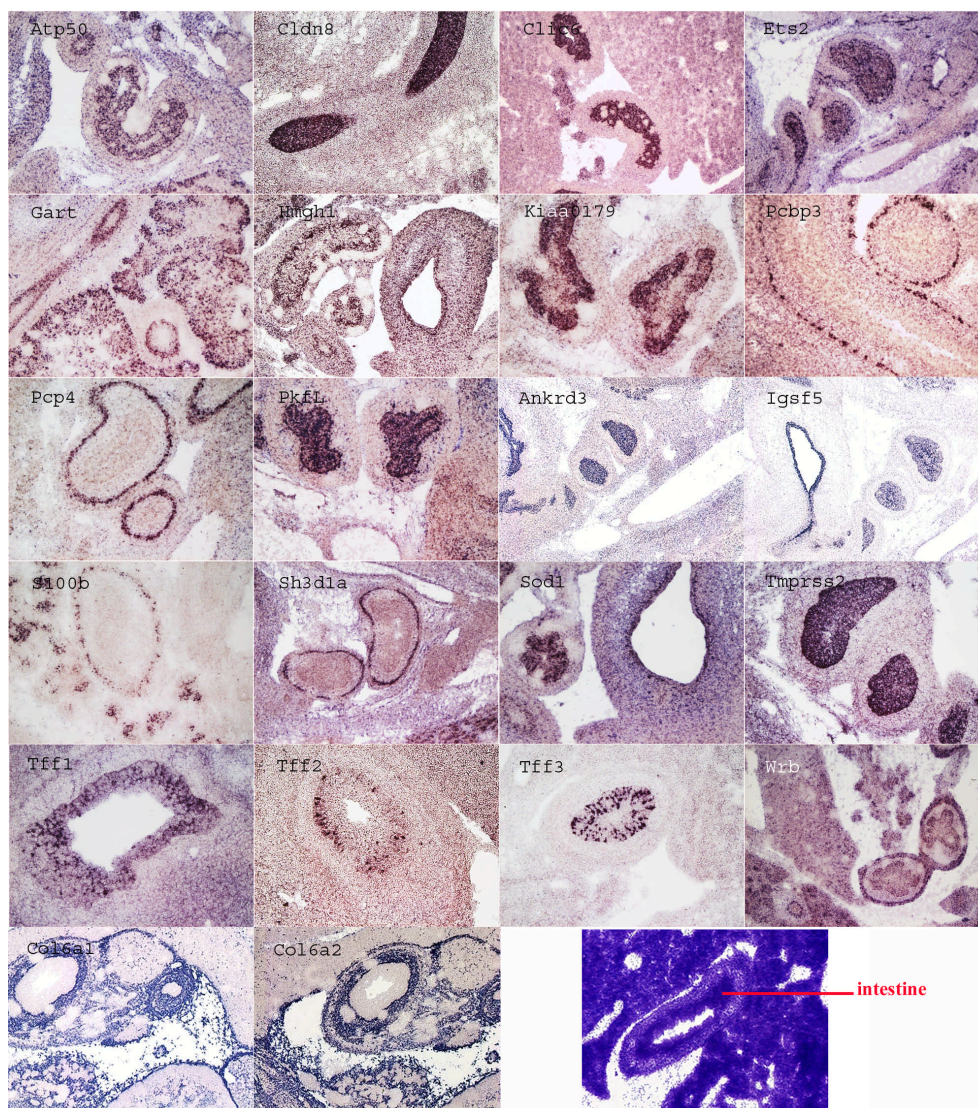


Figure 29. Panel of genes expressed within the gastrointestinal tract. *Atp50*, *Cldn8*, *Ets2*, *Clic6*, *Gart*, *Hmgh1*, *Kiaa0179*, *Pkfl*, *Pkk*, *Igsf5*, *Tmprss2* and *Wrh* were expressed within the endoderm of the duodenum. *Ets2* transcripts were also detected in the vasculature surrounding the duodenum. *Sod1* mRNA was present in the midgut endoderm and in the stomach endothelium. *Tff1*, *Tff2*, *Tff3* were transcribed within a subgroup of cells within the gastro duodenal junction region of the pyloric sphincter. *Pcp4*, *S100b* and *Pcbp3* were expressed within the peripheral

nervous system. *Sh31da* (*Sh3bgr*) was expressed in the smooth muscles of the intestines. *Col6a1* and *Col6a2* show expression within the basal end of the duodenum.

3.2.6 Genes expressed within the thymus

Infection and malignancy are two other major causes of death for DS individuals. The increased risk for infection seems to be caused by abnormalities in the immune system. Compared with normal humans, DS patients show a higher incidence for certain tumors, but they also have a lower risk for other tumor types. This symptom is associated to the thymus which is a component of the immune system, thus in Figure 30 expression in the thymus is presented.

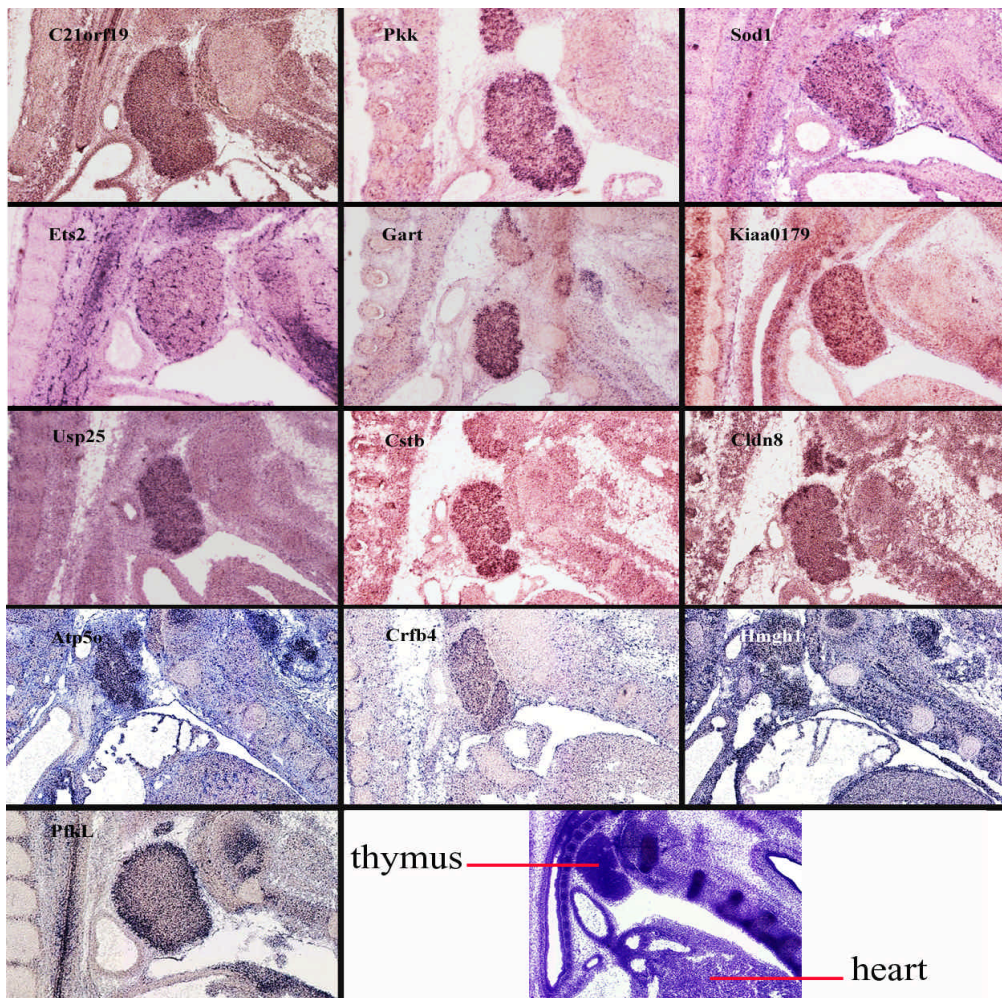


Figure 30. Genes showing expression within the thymus. *Ets2* was expressed within the vasculature of the thymus, all the other genes showed broad ubiquitous expression throughout the thymus.

3.2.7 Genes expressed within the pancreas

Not all but some DS patients have an annular pancreas defect, which occurs when pancreatic tissue encircles the small intestine. This collar around the intestine causes constriction, thus blocking or impairing the flow of food to the rest of the intestines. A panel showing genes expressed in the pancreas at embryonic day 14.5 is presented (Figure 31). At this stage the pancreas is not yet fully formed.

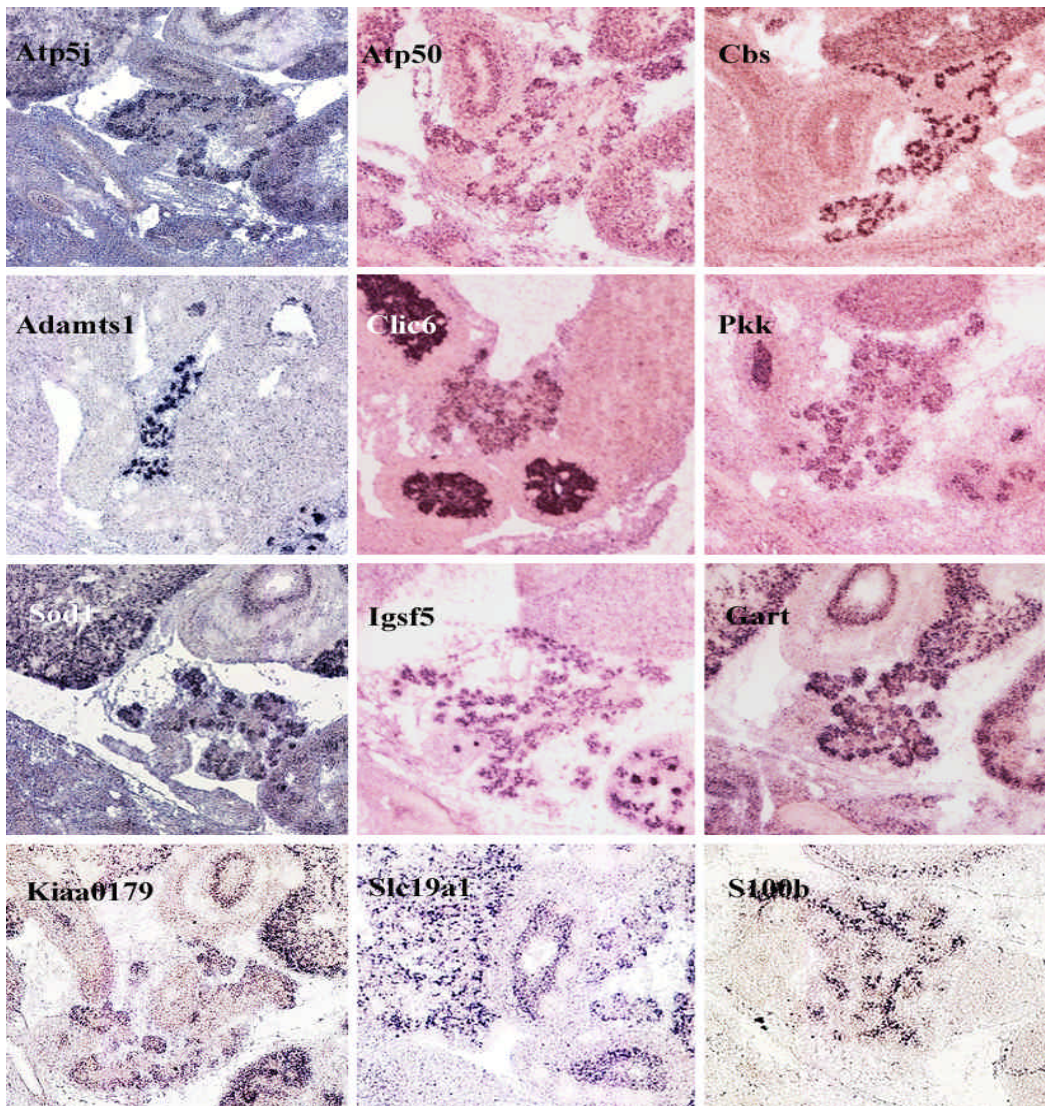


Figure 31. Genes showing expression in the developing pancreas.

3.2.8 Phenotypic comparison of Down with Williams Syndrome

As the expression analysis of the HC21 genes was proceeding, we had the opportunity to analyze the genes in a deletion region linked to Williams syndrome. Williams syndrome occurs from the heterozygous deletion of a ~1.6Mb segment from the human chromosome 7q11.23. Although DS is caused by an extra copy of all the genes on HC21

and Williams syndrome results from the absence of one copy of genes, both these disorders have similar defects (Table 11). Both disorders have their specific phenotypes, too, but there seems to be a correlation between the genome, transcriptome and disease phenotypes, where some genes might be involved in shared regulatory pathways. Table 11 was modified from <http://www.orpha.net> where the mutual features of Down and Williams Syndrome were extracted from a long list of phenotypes.

Table 11. Phenotypic comparison of Down with Williams Syndrome

Down Syndrome	Williams Syndrome
clinodactyly of fifth finger	clinodactyly of fifth finger
congenital cardiac anomaly	complex heart disease (cardiac septal defect)
flat face	flat cheek bones
short/small nose	flattened nose
mental retardation	mental retardation
macrocephaly	microcephaly
microdontia	microdontia
mouth held open	mouth held open
simian crease	simian crease
thick lips	thick lips

We found that some of the DS and Williams syndrome genes had overlapping expression patterns. The mouse orthologue genes of the *intersectin* gene located on HC21 (DS) and the *syntaxin1a* gene located on chromosome 7q11.23 (Williams syndrome) have very similar expression patterns at E14.5. At this critical developmental stage other genes like *elastin* and *C21orf7* also have overlapping gene expression patterns in the outflow tract of the heart. Figure 32 illustrates the similarity between the expression patterns for the above mentioned genes. *Elastin* and *C21orf7* are expressed in the heart where *elastin* has been associated to atrioventricular defects in the heart which are also observed in DS patients. *Intersectin* and *syntaxin1a* are expressed in the brain, and mental retardation is a shared phenotype between Down and Williams syndrome.

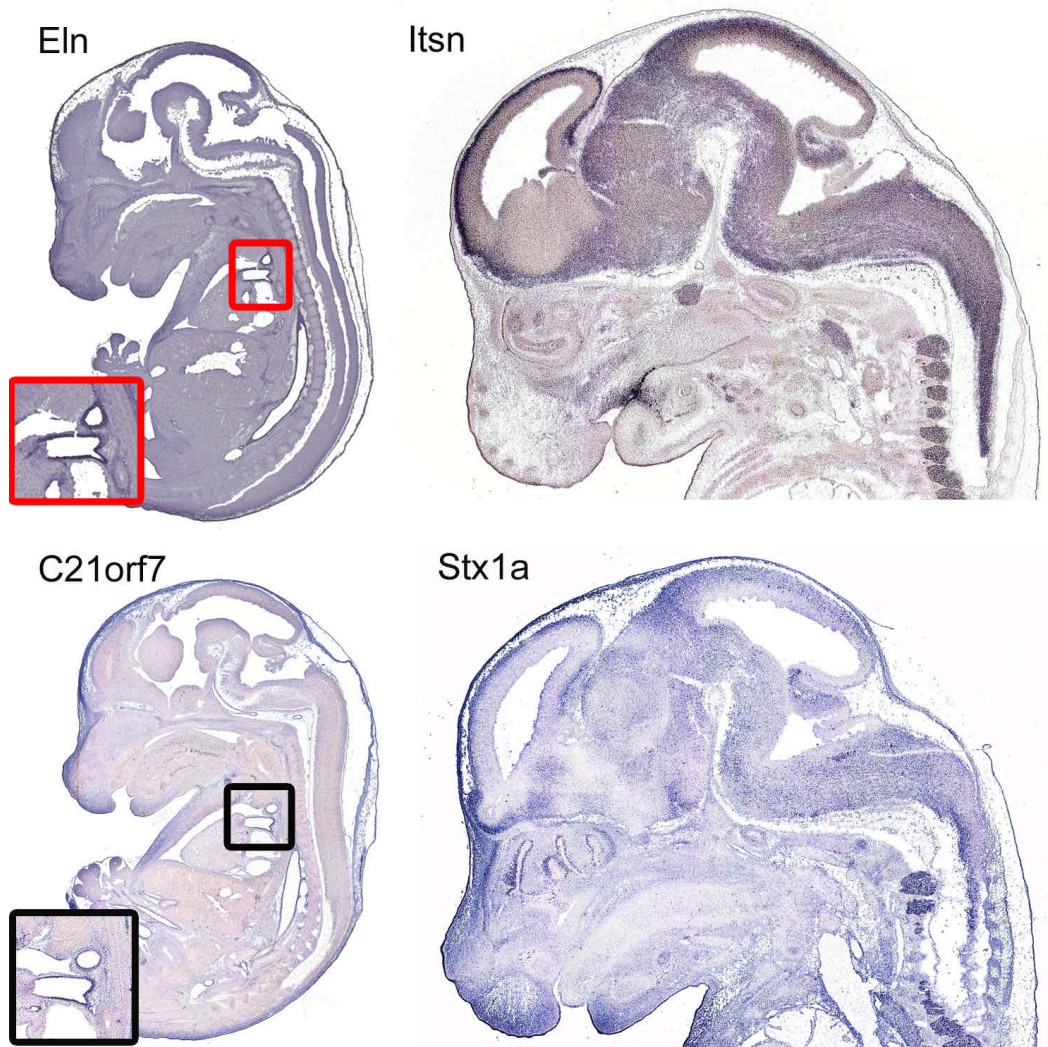


Figure 32. Similarity between the expression patterns for Down and Williams syndrome genes. Both the *elastin* and *C21orf7* gene were expressed in the aortic trunk (red and black box). The expression for the *intersectin* and *syntaxin 1a* gene in the central and peripheral nervous systems was very similar.

3.2.9 Implications of clustering for regulation of expression

The significance of the expression atlas of HC21 genes in relation to genetic disorders is obvious. Illnesses occurring through mishaps that affect one or more genes located on HC21 will eventually be pinpointed, allowing us to prevent or treat these diseases. Based on expression patterns it can easily be seen that the misregulation of the HC21 genes, as in the case of DS, may cause mental retardation. The expression of the genome is orchestrated to permit the proper and natural development of an organism. Through communication carried out across various levels of genomic hierarchy the orchestration of gene expression is achieved with cis and trans regulatory elements. The sequential

arrangement of the genes could provide a mechanism for their coordinated regulation through regulatory elements that steer coordinate transcription of genes, which might be simultaneously required at a certain stage of development. This is what might be occurring for the 9 gene cluster which is illustrated in Figure 33. These 9 genes are on the HC21 syntenic region on mouse chromosome 16. The site of expression of these genes is presented in Figure 34. All 9 genes are expressed in the brain, where 8 of the 9 genes show patterning within the brain. Some of the expression patterns are identical (*Ifngr2*, *Itsn* and *C21orf4*) while some are complementary (*Donson* and *Itsn*). Since the expression patterns of all these genes do not perfectly match, signifying co-expression in the same tissue, it is not easy to suggest a functional relationship.

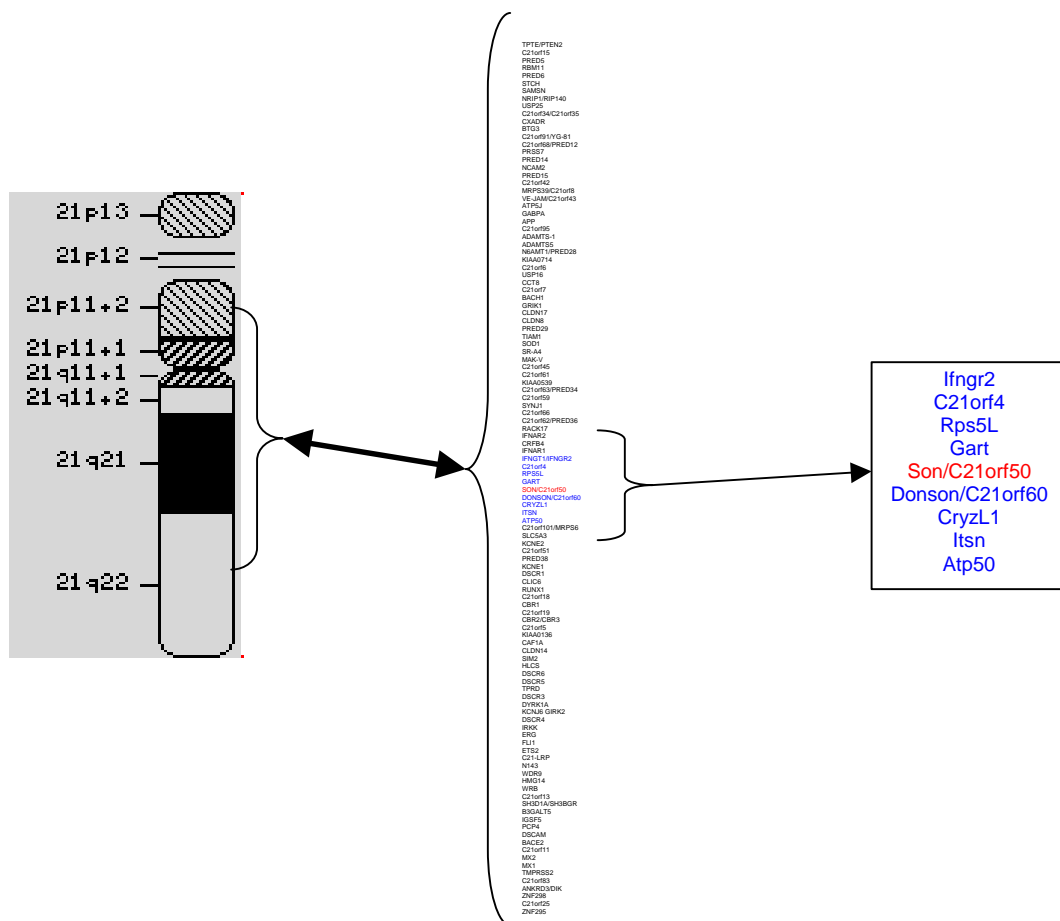


Figure 33. A group of consecutive genes showing expression in the brain. The genes located in the box; *Ifngr2*, *C21orf4*, *Rps5L*, *Gart*, *Donson*, *CryzL1*, *Itsn* and *Atp50* show regional expression in the brain, except for *Son* which is ubiquitously expressed. All genes are on the HC21 syntenic region on mouse chromosome 16.

Gene Name	<i>Ifngr2</i>	<i>C21orf4</i>	<i>Rps5l</i>	<i>Gart</i>	<i>Son</i>	<i>Donson</i>	<i>Cryz11</i>	<i>Itsn</i>	<i>Atp50</i>
Tissue quality	**	**	**	**	**	**	**	**	**
Telencephalon	**R	**R	**R	*R	*U	*R	*R	**R	**R
Diencephalon	0	0	*U	0	*U	0	0	0	*U
Mesencephalon	*R	*R	*U	*R	*U	*U	*R	**U	**U
Rhombencephalon	**R	**R	**R	0	*U	*U	*U	**U	**U
Cerebellum	0	0	*U	0	*U	0	0	0	*U
Spinal Cord	**R	**R	*U	*R	*U	*U	*R	**U	**R
Choroid plexus	0	0	*U	0	*U	0	0	0	*U
Cranial Ganglia	**U	**U	*U	*U	*U	*U	*U	**U	*U
Dorsal Root Ganglia	**U	**U	*U	*U	*U	*U	*U	**U	**U
Parasympathetic ganglia	**U	**U	*U	*U	*U	0	*U	**U	*U

Figure 34. Sites of expression of *Ifngr2*, *C21orf4*, *Rps5L*, *Gart*, *Son*, *Donson*, *Cryz11*, *Itsn* and *Atp50*. These genes are located on the same mouse syntenic block on chromosome 16, and they show the same clustering on the HC21. All genes in the cluster seem to show expression in the peripheral and brain.

Because of the similarity in the pattern of expression we utilized bioinformatics to search for conserved sequences within the 5' region of the cluster genes (up to 20kb). We searched for small (under 100bp) conserved sequences within the same species (intragenomic), for the group of nine consecutive genes (*Ifngr2*, *C21orf4*, *Rps5L*, *Gart*, *Son*, *Donson*, *Cryz11*, *Itsn* and *Atp50*). Among the clustered group of genes 2 conserved sequences were found (the following genes are not 100% conserved among all genes, the percentage drops to ~90%) :

Sequence 1:

```
gagacagggtttctctgtgtagccctagctgtcctggaactcactctgtagaccaggctggactcaactcagaaaaccgctgctctgctc
ccaagtgctgggattaaaggcgatgccaccacaccagctt
```

Sequence 2:

```
gagcggggcgtggtggcgcacgccttaatcccagcactcgggaggcagacgcaggcagatttctgagttcaggccagcctgatctacag
agtgagttccaggacagcagggtacacagaaaacctgtctcg
```

These sequences have ~91% identity to each other at the nucleotide level and are not specific to the cluster genes, they were also found in the 5' region of other genes. A successive cross-species (mouse-rat-human) sequence comparison was carried out for the two sequences. A ~100 nucleotide region is conserved among rat and mouse within the 5' of the cluster of genes with ~80% identity at the nucleotide level. A smaller 49 to 55 base pair region is conserved among human and mouse within the 5 prime of the cluster of

genes with ~90% identity at the nucleotide level. The region showing conservation among mouse, rat and human might therefore be the critical regulatory domain.

To better observe the layout of the cluster genes on the chromosomes and visualize the distribution of sequence 1 and 2 amongst these genes figures were prepared using data provided from ENSEMBL. The mouse, rat and human syntenic regions were determined. From Figures 35-37 the varying length of the syntenic regions can be seen. The distance the exons of the genes are spanning is shown with blue lines, and 5' to 3' orientations are indicated with red arrows. The positions for sequence 1 (green) and sequence 2 (pink) are shown for both mouse (16) and human (21) chromosomes. It can clearly be seen that sequence 1 and 2 are not only in the 5 prime region but they are also placed among the first and last exons. Which is a not surprising for regulatory elements. Some of the conserved sequences were very close to each other, so it was not possible to show each with one arrow on a diagram of such scale. Thus 3 arrow sizes were used which correspond to 1, 2 and 3-4 repeats of sequence. Mouse sequence 1 and sequence 2 have 91% identity at the nucleotide level, and when preparing the rat figure, sequence 2 was localizing to the same nucleotide positions sequence 1 was localizing to, which is why both sequence 1 and 2 are illustrated with one arrow in Figure 37.

Mouse chromosome 16

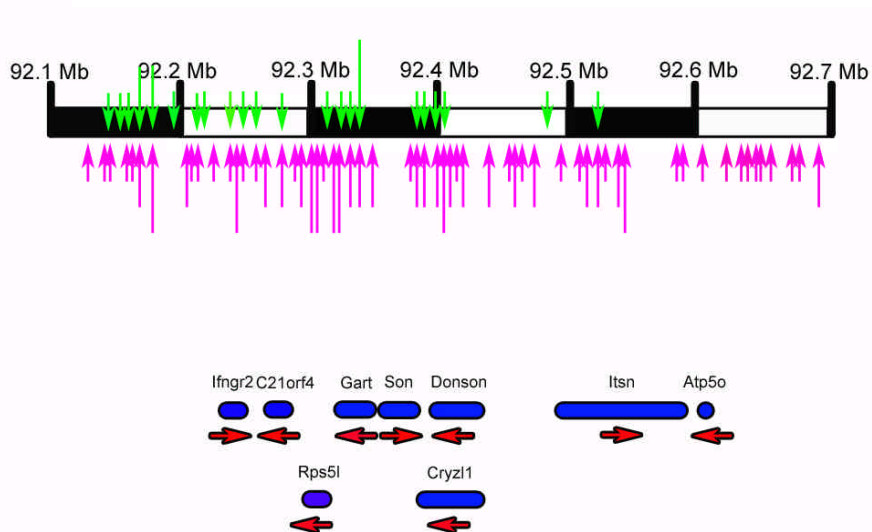


Figure 35. The organization of the clustered genes on mouse chromosome 16 (~0.6 Mb). The cluster genes have 5' to 3' orientations shown with the red arrows. The positions of sequence 1 (green) and sequence 2 (pink) are shown. The 3 arrow sizes correspond to 1, 2 and 3-4 sequence repeats.

Human chromosome 21

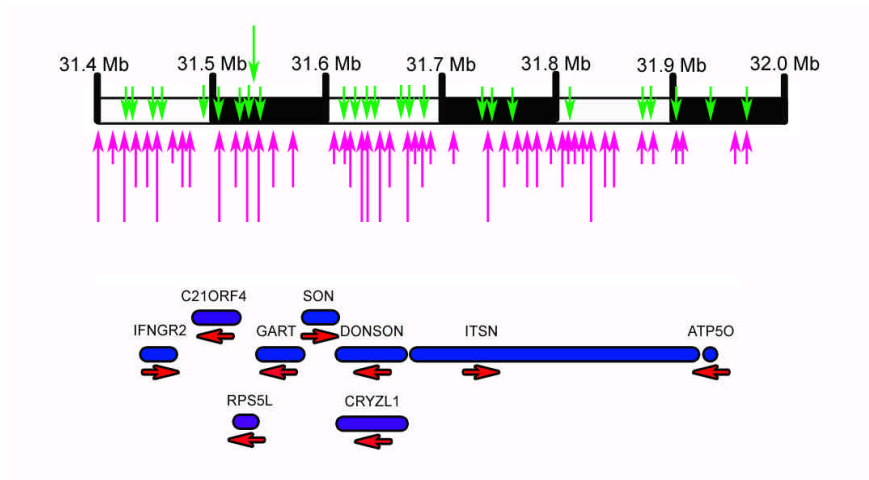


Figure 36. The organization of the clustered genes on human chromosome 21 (~0.6 Mb). The cluster genes have 5' to 3' orientations shown with the red arrows. The positions of sequence 1 (green) and sequence 2 (pink) are shown. The 3 arrow sizes correspond to 1, 2 and 3-4 sequence repeats.

Rat chromosome 11

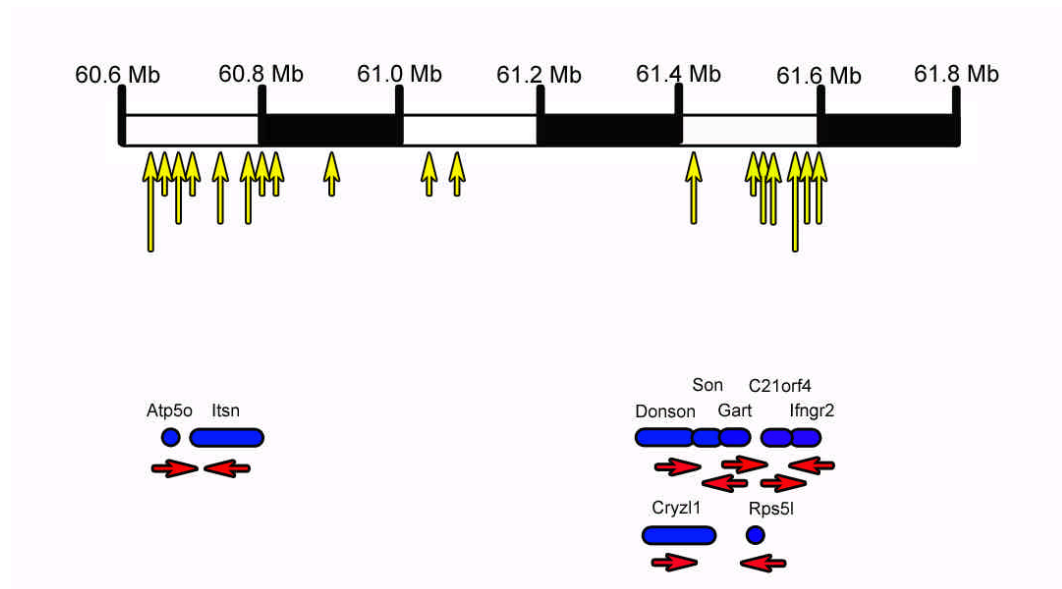


Figure 37. The organization of the clustered genes on rat chromosome 11 (~1.2 Mb). The cluster genes have 5' to 3' orientations shown with the red arrows. The positions of sequence 1 and 2 (yellow) are shown. The 3 arrow sizes correspond to 1, 2 and 3-4 sequence repeats.

Finally a 0.1Mb region from mouse chromosome 1 was used as control (90000kb-90100kb), both sequence 1 and 2 localized to this region. These findings are indicative that the transcription of the genes in the mouse chromosome 16 gene cluster region might be synchronized for the correct development of the brain.

3.3 GenePaint database

The ultimate aim would be to integrate an annotation interface into a database where users will be able query genes expressed at specific stages and tissues. Web-based databases are an excellent means to efficiently retrieve and disseminate scientific data. The large quantities of data from a variety of specimens require a database, which is a powerful tool able to track data acquisition and storage. An interface providing these features has already been developed and recently the annotation interface that will display the annotation of the uploaded genes has been integrated. The annotation of these genes will provide a description of how the expression is patterned (also strength) and at which histological location within the tissues it appears in. Later this will be merged with a search interface that is able to query the annotation (see figure 38) of the stored genes, thus enabling people to search for relevant data (Visel 2002). All mouse orthologues of the human chromosome 21 genes are currently available on the GenePaint.org database (see Appendix 7.2 for more information).

The screenshot shows the GenePaint.org website in a Microsoft Internet Explorer browser window. The address bar displays the URL: http://www.genepaint.org/cgi-bin/mgrqcgi?APPNAME=genepaint&PRGNAME=analysis_viewer&ARGUMENTS=AQ50038739171103,-AMY,-A75,-Asetview. The page features a navigation menu with links for GenePaint Home, Quick guide to GenePaint, Brain Maps, and Advanced Search. An 'Info Box' on the left provides details for the Pcp4 gene, including its accession number (NM_008791), RNA probe (85), orientation (Antisense), hybridization and washing temperatures (59.0), stage (E14.5), strain (NMRI), sectioning plane (sagittal), resolution (5x), and date of dissection (03.Dec.2001). A central image shows a brain section with gene expression patterns. Below the image is a zoom control for 'Embryo_EX76_1_4D'. On the right, an 'Image Directory' table lists various embryo images with their positions.

in Viewlist	Label	Position
<input checked="" type="checkbox"/>	Embryo_EX76_1_1A	0.000
<input checked="" type="checkbox"/>	Embryo_EX76_1_1B	0.120
<input checked="" type="checkbox"/>	Embryo_EX76_1_1C	0.240
<input checked="" type="checkbox"/>	Embryo_EX76_1_1D	0.360
<input checked="" type="checkbox"/>	Embryo_EX76_1_2A	0.480
<input checked="" type="checkbox"/>	Embryo_EX76_1_2B	0.600
<input checked="" type="checkbox"/>	Embryo_EX76_1_2C	0.720
<input checked="" type="checkbox"/>	Embryo_EX76_1_2D	0.840

Figure 38. The GenePaint.org database and the recently integrated annotation interface shown in the lower left corner. The annotation of the *Pcp4* can be seen in the lower left corner. Annotation of the pattern and sites of expression will be accessible through a query interface.

To sum things up, the usefulness of the automated non-radioactive RNA ISH technology developed in our lab was exemplified. First the TSA protocol that improves the sensitivity was examined, and the gain in signal strength in presence of the TSA was shown. The low background of the non-radioactive method was also illustrated by comparing expression in wildtype and mutant knock-out mice, where very low background was obtained in mutant knock-out mice when the tissue was probed for the transcript that was knocked out. Second, to show that the automated non-radioactive RNA ISH protocol used in this study was within the range of sensitivity that is attributed to conventional (^{35}S -UTP) radioactive RNA ISH methods, results obtained with radioactive procedures were compared with non-radioactive TSA enhanced RNA ISH. Generally there was excellent agreement between the results. Finally, using a weak but uniformly expressed gene (*Dscr3*), the sensitivity, in terms of mRNA copy number per cell of the non-radioactive RNA ISH was quantified by measuring the number of transcripts present in a defined tissue volume by PCR. As expression was detected with as few as 3 molecules of *Dscr3* mRNA, the non-radioactive ISH procedure used for the present work was found to be very sensitive, similar to the sensitivity reported for ^{35}S based ISH (Speel et al. 1998).

After demonstrating that the non-radioactive ISH protocol was optimized to a degree where it was applicable to a biological question, the expression of the mouse orthologues of the human chromosome 21 genes was determined. This was the first large-scale application of the automated non-radioactive ISH technology. The objective here was to gain functional information by localizing the transcript to tissues and organs. Since the DS physical phenotypes have already been associated to organs like the brain (mental retardation), heart (congenital heart defects), skeleton (dysmorphic features), gut (Hirschspung disease), thymus (infection and childhood leukemia) and pancreas (annular pancreas), panels showing genes expressed in these organs were prepared. Additionally a web page (www.tigem.it/ch21exp) was formulated and put up for the easy access of images, annotation tables and experimental metadata for the “human chromosome 21 gene expression atlas”.

IV. Discussion

Analysis of the sequence of the complete human and mouse genomes has provided a first look at the approximate gene content (Lander et al. 2001, Mouse Genome Consortium 2002). This progress gave rise to a new challenge which is the large scale analysis of gene function. Visualizing gene expression within specific cells of known function assists in unravelling gene function. The instrumentation to carry out genome wide analysis of gene expression using an automated RNA ISH protocol was developed and tested. This technology has the capability to determine and represent the expression patterns of thousands of genes within an acceptable time scale. After demonstrating the accuracy and reliability of the automated non-radioactive RNA ISH technology, we analyzed, as an application, the ~200 mouse orthologues of HC21 genes, to reveal candidate genes associated to the various DS phenotypes. The expression pattern data not only implicated an association between candidate genes and disease symptoms, but provided evidence for the existence of DNA control regions in chromosomes that could control gene expression.

4.1 Evaluation of the non-radioactive *in situ* hybridization protocol

It is believed that non-radioactive ISH is less sensitive than radioactive RNA ISH, yet advances like a catalyzed reporter deposition (CARD) based tyramide signal amplification technique that can be used in the non-radioactive ISH protocols, has provided a major improvement in sensitivity (Bobrow et al. 1989). Methods like Northern blot and RNase protection assay rely on RNA extracted from bulk tissue. This may result in mRNA transcripts expressed within a few specific cells to be highly diluted making them appear rare (or non-detectable) when they are alone among many cells that do not express this transcript at all. When radiolabeled probes are used for ISH, it can be difficult to obtain the resolution required to identify individual cells because the signal consists of silver grains in an autoradiographic emulsion overlaying the tissue section. By using three independent approaches the accuracy of the automated non-radioactive RNA ISH technology developed in our lab was shown.

The TSA step which provides a crucial increase to the sensitivity of the non-radioactive ISH was explored. The gain in signal strength in the absence and presence of the TSA was examined. This was done by probing two specimens, where when the transcripts in one set of specimens were being detected, the protocol including the tyramide signal amplification step was used. This step was left out for the second set of specimens where

the *Pcp4* gene transcripts were probed in both specimens. In the specimen which was probed omitting the TSA enhancement step the signal nearly totally disappeared. It was even very hard to see the tissue itself where the tissue looked as if it was not hybridized with RNA probe and signal was very faint, only being observed in regions of strong expression. In the protocol provided by the manufacturer it is pointed out that TSA may result in a high background which could be because of endogenous biotin or enzymes (alkaline phosphatase and peroxidases) that were not suitably inactivated. The automated system allows the use of multiple pipetting steps with accurate reproducibility, thus with stringent washing steps and the treatment with blocking solutions containing inhibitors, a decrease artifact background is obtained. The low background of the non-radioactive method was explored by comparing wildtype and null mutant mice, where very low background and no signal was obtained as expected when the tissue was probed for the transcript that was knocked out. This was done with the *Ear2* and *Prkcb* mutant mice. The absence of the transcript in these knock out mice had previously been verified with Northern blots (personal communication, Leitges, Warnecke, 2003).

Although the *Prkcb* mutant tissue specimen was expected to be blank, the first probe we designed for this assay gave a significant signal in the hippocampus due to cross-reactivity with transcripts of other PKC isoforms. Analysis of the probe sequence revealed a 700 nucleotide long region that was conserved between *Prkcb* other isoforms of the PKC family of genes. When this region was removed the cross-reactivity disappeared and sections of *Prkcb* mutant brains were blank. This case shows that the protocol works well but one needs to ensure a gene specific template (sequence content) design.

A second approach that was taken aimed to show that the automated non-radioactive RNA ISH protocol used in this study, compared favorably with the range of sensitivity attributed to conventional (^{35}S -UTP) radioactive RNA ISH methods. There was excellent agreement between the patterns of expression when data created by the two methods were directly compared. This comparison was done with genes exhibiting a variety of expression patterns. Since the non-radioactive method is faster (approximately 24 hours from beginning to the end) than the radioactive ISH, has cellular resolution, and does not require work with dangerous radioactive compounds, radioactive methods seem to lose their attraction. The non-radioactive ISH protocol used in this study also allows the easy evaluation of the tissue histology, and an additional histochemical staining step required in the case of the radioactive method are no longer necessary. When one uses silver

grains to pick up the radioactive signal and Hoechst to stain the nucleus of the cells the tissue has to be photographed with two microscopic settings (darkfield and fluorescence) followed by merging the resulting images with image software. This is done by adjusting the images to each other with software which decreases the objectivity of the data. The precipitate and the histological staining resulting from non-radioactive ISH protocol is observable under one microscopic setting (brightfield).

To directly examine the sensitivity of the non-radioactive ISH protocol used in the present study, it was necessary to quantify the detectable transcript copy numbers to determine if the sensitivity was similar to that reported for ^{35}S based ISH (Speel et al. 1998). Using quantitative PCR, the sensitivity (i.e. copy number per cell) of the non-radioactive RNA ISH was estimated by measuring the number of transcripts present in a defined tissue volume for a weak and uniformly expressed gene (*Dscr3*). The total DNA content of the tissue was used to estimate the number of cells. We found that hybridization signal could be seen if as few as 3 molecules of *Dscr3* mRNA per cell were present which is the same range as the sensitivity reported for ^{35}S based ISH (Speel et al. 1998). One point to keep in mind is that the tissue specimens used in this copy number assay were not pretreated like the sections that are used for the typical ISH protocol. In the typical protocol, tissue sections have to go through a series of pretreatments (fixation, acetylation, dehydration, peroxide treatment, HCl, proteinase k digestion) before actually being hybridized with the hapten labeled probe. These treatments probably cause the tissue to lose RNA transcripts. This would mean that the procedure detects less than what is present in the tissue used in the PCR quantification experiments, implying that the ISH procedure would be even more sensitive than what we have estimated. On the other hand a difficult to account loss might be occurring when extracting the mRNA from the tissue specimen, which would imply that we over estimate the sensitivity.

Asan and Kugler (1995), Zreiqat and coworkers (1998) have shown that there is a linear relationship between the amount of alkaline phosphatase and the amount of mRNA detected in non-radioactive ISH. Together with the systematic investigations that were carried out, the current automated non-radioactive ISH method is both qualitative and quantitative. After assuring the readiness of the technology, the final test to further prove this point was carried out, the technology was applied to the mouse orthologues of the HC21 genes.

4.1.1 Quality of *in situ* hybridization data depends on the source of the DNA template

Different IMAGE, BMAP-EST, NIA 15K mouse cDNA and NIID clones were ordered and used in this study to generate hapten labeled RNA probes. To actually have such an easily accessible cDNA resource representing the whole genome is of course an advantage when carrying out genome-wide studies. There are however, two faces to the coin. When working with such a vast number of clones, some get mixed up by the suppliers that create and provide these clones especially during the rearranging process. Hence it was necessary to sequence verify that the clones at hand were actually those that were ordered. To obtain a gene specific template, by reverse transcription followed by subcloning and sequencing, is time consuming. Moreover, the resulting clones will have to be made linear for subsequent *in vitro* transcription. The whole process was streamlined as follows;

1. The T7 and SP6 promotor sequence was added to the gene specific primer sequence used for RT-PCR removing the requirement to actually subclone the cDNA template into a vector. This eliminates the need to work with bacteria, as is required with the cDNA template subcloned into a vector plasmid. With a PCR approach it is possible to store and later to reamplify the template of interest. This would also imply that the step where the plasmid DNA is to be linearized with restriction enzymes before transcribing the RNA probe, is no longer needed.
2. Another advantage of this strategy was that since the sequence of the whole mouse and human genomes are available, one can use bioinformatics to design templates covering specific regions of the gene of interest.
3. When designing such templates the poly A tails that can create high background as a result of non-specific binding can be left out.
4. The probe length which is most optimal under specific experimental conditions can be chosen. Probe length is important because a too short probe might not give a detectable signal. The longer probes contain more nucleotides, thus they have more dig-UTPs, resulting in a stronger signal.

For the HC21 orthologue mouse genes *IFNGR2*, *RUNX1*, *DYRK1A*, *GRIK1*, and *JAM2* new probes were redesigned based upon the above PCR strategy. These were chosen among 15 genes that had a questionable expression pattern, as they were either inconsistent with published data or had a high background-like signal. The improvement in the results was remarkable. For example the new *Jam2* template gave very specific

signal. At the structural level the new probe had no poly A tail and encompassed a more 5' region of the mRNA.

4.2 Analysis of the HC21 orthologue mouse gene expression data

By combining gene mapping with expression analysis it is possible to identify candidate genes for human diseases (Ballabio 1993). The expression study of the HC21 mouse orthologues thus provides a rich resource for candidate genes for both monogenic and multigenic diseases mapping to HC21. However, the greatest impact of this study lies in assessing the contribution of specific genes to DS traits and phenotypes. DS affects 1 out of 700 live births and is caused by a third copy of the 21st human chromosome which results in severe mental retardation and dysmorphogenesis of various organs (Wiltshire et al. 1999). After a 99.7% sequence coverage of chromosome 21q was achieved revealing 178 confirmed genes and 36 predicted genes (Hattori et al. 2000) the opportunity to gain an understanding of the pathogenetic mechanisms leading to DS was revealed. To gain a functional perspective of the processes in which the known, novel and predicted genes on HC21 were involved, a joint effort was put into creating a high resolution expression “atlas” of human chromosome 21 (Reymond et al. 2002b).

The mouse orthologues of the HC21 genes were analyzed via automated non-radioactive ISH to gain functional information by localizing the transcript, to a tissue with a specific function (e.g. pancreas, limbs). The fact that a gene is expressed at a particular stage of embryonic development in a specific tissue or organ, would indicate that this gene may have a function in that tissue or an important role in developmental regulation (Reeves 2002).

Expression was detected for 64% of the tested genes examined at embryonic day 14.5. Regional gene expression was observed in 42% of the genes. Ubiquitous expression was observed in 13% of the cases. Genes with both weak ubiquitous and strong regional expression were observed in 9% of the cases. No expression was detected for 36% of genes. Tissue distribution of regionally expressed transcripts reflects an association with the DS phenotype. The largest portion of genes showing regional expression was within the brain. It is only logical to predict that their misregulation (as in the case of trisomy 21) would lead to a phenotype associated to the brain. This positive correlation is followed by a number of genes expressed in other organs affected in Down syndrome such as the gut. Organs like the limbs, heart and thymus also express genes in a regional manner. The smallest portion of regionally expressed genes are found in the pancreas. The DS physical

phenotypes have been associated to the following organs: the mental retardation phenotype where DS patients have a lower IQ when compared to normal individuals and Alzheimer's disease has been associated to the brain. The DS heart defects where atrioventricular defects in 39% of the children, and congenital heart disease is observed in 40% of the adults, has been linked to the heart. The DS dysmorphic features like abnormal limbs, fingers and toes, arise due to defects in the skeleton. Intestinal lesions observed among Down patients have been linked to the intestines. There are indications that DS individuals have an altered immune system due to different risk factors for various tumors compared to normal individuals, and because they are susceptible to infection. For this reason, although it is not a primary critical DS organ, the immune defects have been linked to the thymus. The annular pancreas defect observed in DS has been associated to the pancreas.

4.2.1 Down syndrome

4.2.1.1 DS brain associated defects

Mental retardation is a feature of DS which reflects the most complicated and limiting of its handicaps. Until now no pharmacological therapy has been shown to have a positive effect. Although a decreased density of neurons, changes in phospholipid composition and alterations in electrophysiological properties of the brain and isolated neurons have been studied, the specific causes for mental retardation are still not known (Epstein 1995). The numerous HC21 orthologue genes expressed in the brain, increase the risk of a phenotype occurring in this organ due to improper development caused by misregulation. The genes expressed at this critical time point (embryonic day 14.5) are important for the proper development of the brain. There are transgenic mice that have been associated with DS. The *Sod1* gene shows ubiquitous expression with strong regional expression in the brain and spinal cord, and *Sod1* transgenic animals show learning defects (Epstein 1995). Axonal proliferation and astrocytosis is observed in the *S100 β* transgenic mice (Reeves et al. 1994).

Another defect observed in the brain of Down patients is that they show the pathological changes observed in people with Alzheimer disease (Busciglio et al. 2002, Frederikse and Ren 2002, Head et al. 2002, Motonaga et al. 2002). Although Alzheimer is usually observed at old age in normal individuals, DS patients show the disease symptoms after their third decade. Early onset Alzheimer disease is a very important disease related to neuronal degeneration occurring through the accumulation of amyloid plaques. Like the amyloid precursor protein gene (*APP*) which has already been associated to early onset

Alzheimer disease, other HC21 genes showing regional expression in the brain like *SYNJ1*, *S100B*, *NRIP1*, *TTC3*, *CCT8*, *GRIK1*, *KCNJ6*, *IFNGR2*, *HUNK* and *PCP4* are candidates for further studies to solve this disease. Beta site amyloid precursor protein cleaving enzyme (*BACE2*) is another HC21 gene which has been linked to Alzheimer which shows expression in the brain (Benett et al. 2000). Understanding Alzheimer might even help us gain insight into other common diseases that occur through neuronal degeneration such as Parkinson disease. All the above mentioned genes are also expressed in the midbrain-hindbrain junction which has been associated to dopaminergic neurons, which degenerate in Parkinson disease patients.

4.2.1.2 DS heart associated defects

One of the major causes for death among people with DS is congenital heart disease. Congenital heart disease is observed in 40% of people with DS. 39% of the children affected by DS have been reported to have atrioventricular defects among other cardiac lesions (Epstein 1995). In particular, *ADARBI*, *KCNJ15*, *PFKL*, *C21ORF7*, *ETS2* and *SH3BGR* are promising candidate genes for the DS heart abnormalities because their murine orthologues are expressed in the developing heart. Apart from congenital heart disease, HC21 has also been associated to other heart diseases like Long QT syndrome and Bethlem myopathy, and thus provides an important model to link individual genes to pathways controlling heart development. The genes of HC21 are expressed in particular regions of the heart during embryonic development; therefore these regionally expressed genes may have a role in heart defects found in Down patients. Patients with the following heart disease have mutations in the following genes: Bethlem myopathy patients have mutations in *COL6A1*, Ullrich disease patients have mutations in *COL6A2* and Knobloch syndrome patients have mutations in *COL18A1* (Abbott et al. 1999, Sertie et al. 2000). As stated in literature, *Col18a1* was detected in both cardiac blood vessels and blood vessels of the whole body. *Col6a1* and *Col6a2* were strongly expressed in the mitral valve and along the pericardium as already has been described (Camacho Vanegas et al. 2001, Higuchi et al. 2001, Jobsis et al. 1996).

4.2.1.3 DS skeletal defects

An easily recognizable, albeit a less serious DS phenotype, is that all patients have dysmorphic features. DS fetuses exhibit a reduced growth rate of limb long bones during the third trimester of pregnancy. Limb defects such as a gap between the second and first toe, short, broad hands and a non-ossified or hypoplastic middle phalanx of the fifth digit has been observed in these fetuses (Epstein 1995). Several genes of HC21 were

regionally expressed in the limbs at embryonic day 14.5. *Erg* has strong expression in the posterior proximal mesoderm and in the joints. *Ets2* was also expressed in the skeletal system and was previously found to result in skeletal abnormalities reminiscent of DS when overexpressed in transgenic mice (Sumarsono et al. 1996). *Adamts1* was expressed in the perichondrium of the developing bones. Although the limb phenotype is as striking as that of mental retardation, only 10% of the genes examined showed regional expression in the limbs. One possibility is that the few genes regionally expressed in the limbs are involved in a network which does not allow misregulation.

4.2.1.4 DS gastrointestinal abnormalities

DS patients have a high frequency for gastrointestinal abnormalities such as duodenal stenosis, Hirschsprung's disease, gastroesophageal reflux and imperforate anus (Epstein 1995). The digestive tract is well differentiated at embryonic day 14.5 and the expression of genes can be detected easily. The mouse genes *Atp50*, *Cldn8*, *Clic6*, and *Ets2* were expressed within the endothelium of the duodenum. *Tff3* and *Sod1* transcripts were present in a subgroup of cells within the gastro duodenal junction region of the pyloric sphincter, with some expression in the endothelium of the stomach and the intraperitoneal portion of the midgut, respectively. The human orthologues of the gut-expressed genes *ATP50*, *CLDN8*, *CLIC6*, *ETS2*, *KCNJ15*, *SH3BGR*, and *WRB*, may thus play a role in the DS gastrointestinal abnormalities (Kola and Herzog 1997).

4.2.1.5 DS immune system abnormalities

Infection and malignancy are two other major causes of death for DS people (Epstein 1995). These defects seem to arise through abnormalities in the immune system. Down patients show a lower incidence of certain tumor types: neuroblastomas, and solid tumors, a higher risk for other tumor types; leukemias, lymphomas, testicular germ cells (Satge et al. 1998a, Hasle et al. 2000). The decreased incidence of solid tumors in individuals with DS, indicates that the increased dosage of HC21 genes may be protecting these individuals from these tumors (Satge et al. 1998b, Hasle et al. 2000). Satge et al. have associated the overexpression of *SI00B* in Down patients with the lower risk for neuroblastomas. Loss of heterozygosity has been observed for specific regions of HC21 in several solid tumors (Sakata et al. 1997, Kohno et al. 1998, Ohgaki et al. 1998, Yamamoto et al. 1999, Ghadimi et al. 1999, Bockmuhl 1998) including cancers of the head, neck, breast, pancreas, mouth, stomach, oesophagus, and lung. This usually indicates that there is at least one tumor suppressor gene in this region. Recently it has been shown that AML1-ETO can suppress transcription of the p14^{arf} tumor suppressor

gene, through suppression of the *AML1 (RUNX1)* DNA binding site of the p14^{arf} promoter (Linggi et al. 2002).

4.2.1.6 DS pancreas defects

Not all, but some DS patients have an annular pancreas defect. This defect arises when the pancreas tissue forms a ring-like collar around the intestine, preventing the further passage of food into the digestive tract. At this stage the pancreas is still a primordia and not well structured, thus a misregulation of the genes expressed at embryonic day 14.5 could cause the dysmorphogenesis of the pancreas.

4.2.2 Phenotypic comparison of Down with Williams Syndrome

Williams syndrome occurs from the heterozygous deletion of a ~1.6Mb segment from the human chromosome 7q11.23. This corresponds to a 1.4Mb region in the mouse chromosome 5 which has been sequenced with high accuracy (DeSilva et al. 2002). Although DS is characterized by an extra copy of all the genes on HC21, and Williams syndrome is characterized by the absence of one copy of genes, both these disorders have similar defects. A few of the genes on both chromosomes have exhibited overlapping expression patterns. For example, the mouse orthologues of the *Intersectin* gene, located on HC21 (DS), and the *Syntaxin1a* gene, located on chromosome 7q11.23 (Williams syndrome), have identical expression patterns at E14.5. Both these genes have been associated with clathrin mediated vesicle trafficking (Pucharos et al. 2001, Metzler et al. 2001). This data raises the possibility that there is a correlation between the transcriptome and disease phenotypes.

4.2.3 Implications for clustering of expression patterns, for shared regulation

The components of the genome are orchestrated to permit the proper and natural development of an organism. To achieve this, the genome needs to synchronize and influence its components through communication carried out across various levels of hierarchy. This could be done through the co-regulation of genes belonging to shared regulatory pathways owing to shared non-coding sequence motifs that direct the binding of specific groups of shared transcription factors (Li et al. 1999, Dermitzakis et al. 2002). However there is the probability that the co-regulation of genes might be in response to Locus Control Regions (LCRs). LCRs are novel elements, distinct from promoters, classical enhancers, chromosomal insulator, matrix attachment regions or scaffold attachment regions, they open the nucleosome structure so other structures can bind. They are copy number dependant where they temporally influence expression of linked genes

(Schreiber and Bernstein 2002, Pennacchio and Rubin 2001). The consecutive placement of the genes on the chromosome could be to allow their coordinated regulation through an element influencing their synchronized transcription which might be simultaneously required at a certain stage of development.

Our attention was drawn to a group of nine genes (*Ifngr2*, *C21orf4*, *Rps5L*, *Gart*, *Son*, *Donson*, *Cryz11*, *Itsn* and *Atp50*) showing similar expression patterns in the brain. These genes are consecutive both in the mouse and in the human chromosomes. Among these nine genes some have expression patterns which overlap or are identical (*Ifngr2*, *C21orf4* and *Itsn*), while some are complimentary (*Itsn* and *Donson*). All nine mouse orthologues are located on the same mouse syntenic block on chromosome 16, and they show the same order as on the HC21. Since the expression patterns of all the genes do not overlap, it is not easy to suggest a functional relationship between these functionally divergent proteins. This does not exclude the possibility that they could be involved in a common pathway.

We utilized bioinformatics to search for conserved sequences within the 5' region of the cluster genes (~20kb 5' region). We searched for small (under 100bp) conserved sequences within the same species (intragenomic). In a study the Levine lab in Berkley (USA) showed that there are shared regulatory elements involved in the dorsal-ventral patterning (Stathopoulos et al. 2002). We found 2 shared sequences among our clustered group of genes. After the conserved sequences were identified the region of the chromosome containing all the nine genes of the cluster were scanned for these sequences. The sequence was not only found in the 5' region, but also the introns and the 3' regions of the group of nine genes it could be influencing.

A cross-species sequence comparison was also carried out for the two sequences and they were also found to be conserved among rat, human and mouse to a certain degree. The region showing conservation among mouse, rat and human might be the critical regulatory domain. However, functional experiments (Boffelli et al. 2003) which might allow us to determine if these are truly functional regulatory elements have not yet been done. Nevertheless all these findings are indicative that the transcription of the genes in this region might be synchronized for the correct development of the brain. Albeit preliminary, this data suggest that regions containing coexpressed genes exist on HC21. Further studies are required to verify this, but attention needs to be drawn to an extent

that a preliminary overview acquired from genome-wide studies will allow us to obtain the initial data for such assays.

4.3 Future directions

This study does not solve the problems that arise from the presence of a third copy of a whole chromosome. The correlation of the HC21 gene expression patterns, and the disease phenotypes, however, allows us to prioritize various candidate genes, which of course can be further examined with follow up experiments. Appropriate, functional experimental approaches will be developed, based on the type of genes that are prioritized according to expression information gained from this study. For example, since we have already catalogued the native expression pattern of the HC21 mouse orthologue genes, observing the changes occurring in the mutant mice (partial trisomy 16 mouse model) would help us understand what happens under the influence of an extra copy of the HC21 orthologue mouse genes (Sago et al. 1998).

The mouse is a suitable model organism for the study of human biology because it has a well characterized genome, has a similar physiology, it is easy to collect specimens and compared to human the chief difference is mainly in size and shape. Nevertheless it should not be forgotten that trisomy 21 is a unique human disorder. To obtain conclusive data of greater medical relevance this study might need to be followed up with experiments using human embryo specimens. Human tissue samples are rare and valuable, prioritizing genes through a preliminary study carried out in mouse samples makes the whole process relatively feasible. Once these genes are verified in human tissues they would be further prioritized e.g. as drug targets. Thus the next step would be to carry out a bioinformatic analysis to determine the protein sequence, which would reveal if the gene had domains identical to the already existing characteristics of proteins that can be targeted by drugs. These would include channel proteins, G-protein coupled receptors, secreted proteins and kinases among other proteins. With follow-up functional studies, the biomedical relevance (diagnostics, intervention and drug therapy) of these genes would be determined.

By itself the GenePaint database will have the well catalogued and annotated gene expression patterns of the whole genome, which is a striking achievement all by itself. However, in the future query interfaces where one can retrieve the data according to ones interest, (structure, sequence, expression clustering, etc.) will be available. It will be possible when working with specific diseases such as Parkinson's disease to retrieve all

the genes expressed within the substantia nigra and hence to obtain a list of putatively interacting gene products. Since the goal is to carry this out across certain developmental stages, the data will have developmental relevance and shall be even more informative.

The organization and linking of the data to other databases will make it even more powerful. Now the database is linked to GenBank and GeneCards which provide substantial background information to the gene of interest. A future goal will be to search the database for regulatory networks that could shed light on the organization of the genome. Insight into issues like gene clustering could be gained by linking gene expression patterns to chromosomal location or if it were possible to identify shared regulatory elements through various bioinformatics interfaces. Clearly, much work needs to be done for such a large-scale database on gene expression but the present study shows that this effort is both possible and highly promising.

V. Conclusion

The ISH technology (GenePaint) is an integrated system which has been developed to carry out, catalogue and store and retrieve gene expression data. Not only has the automated process of treating the slides rapidly via *in situ* hybridization been increased to a level where a large amount of data can be produced, but the cover slipping, scanning, data cataloguing and image retrieval interfaces (along with automated annotating software) need to be running like the wheels of a smoothly ticking clock. I have tried to present, what could be called, the part of the non-radioactive ISH technology which is directly related to the wet lab experiments. Through the use of 3 approaches the reliability of this automated non-radioactive ISH technology was explored; 1. assessing the advantages of the TSA, 2. comparing with radioactive ISH methods and 3. quantifying the detection limit.

After demonstrating that the non-radioactive ISH protocol was optimized to a degree where it was applicable to a biological question, the expression of the mouse orthologues of the human chromosome 21 genes was determined. This was the first large-scale application of the automated non-radioactive ISH technology. The objective here was to gain functional information by localizing the transcript to tissues and organs. With the analysis of nearly all HC21 genes, another level of genome annotation has been reached. Now it will be easier to pinpoint candidate genes for monogenic disorders and complex diseases which could be used to diagnose the illness in its early stages, or even better, to find a way to prevent the disease from occurring by intervening with its development e.g. via gene therapy. In other words this study is another advance of research into DS and HC21 associated disorders that helps in the search for genes of therapeutic relevance.

Last but not least, this study has provided initial insights for clustering of expression patterns, through shared regulation.

VI. References

Abbott GW, Sesti F, Splawski I, Buck ME, Lehmann MH, Timothy KW, Keating MT, Goldstein SA. MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell*. 1999 Apr 16; 97(2):175-87.

Adams JC. Biotin amplification of biotin and horseradish peroxidase signals in histochemical stains. *J Histochem Cytochem*. 1992 Oct;40(10):1457-63.

Antonarakis SE. Chromosome 21: from sequence to applications. *Curr Opin Genet Dev*. 2001 Jun;11(3):241-6.

Asan E, Kugler P. Qualitative and quantitative detection of alkaline phosphatase coupled to an oligonucleotide probe for somatostatin mRNA after in situ hybridization using unfixed rat brain tissue. *Histochem Cell Biol*. 1995 Jun;103(6):463-71.

Awgulewitsch A, Bieberich C, Bogarad L, Shashikant C, Ruddle FH. Structural analysis of the Hox-3.1 transcription unit and the Hox-3.2--Hox-3.1 intergenic region. *Proc Natl Acad Sci U S A* 1990 Aug;87(16):6428-32

Ballabio, A. The rise and fall of positional cloning? *Nat Genet* 1993 Apr;3(4):277-9

Beaudet AL, Scriver, C R, Beaudet AL, Sly WS, & Valle D. in *The Metabolic and Molecular Bases of Inherited Disease* (eds. Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 3-128 (McGraw Hill, New York, 1995).

Benett BD, Babu-Khan S, Leoloff R, Louis J, Curran E, Citron M, Vasser R. Expression analysis of BACE2 in Brain and peripheral tissues. *J Biol Chem*. 2000 Jul 7;275(27):20647-20651.

Bobrow MN, Harris TD, Shughnessy KJ, Litt GJ. Catalyzed reporter deposition, a novel method of signal amplification. Application to immunoassays. *J Immunol Methods* 1989 125:279-285.

Bockmuhl U, Wolf G, Schmidt S, Schwendel A, Jahnke V, Dietel M, Petersen I. Genomic alterations associated with malignancy in head and neck cancer. *Head Neck*. 1998 Mar; 20(2):145-51.

Boffelli D, McAuliffe J, Ovcharenko D, Lewis KD, Ovcharenko I, Pachter L, Rubin EM. Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science*. 2003 Feb 28; 299:1391-1394.

Boncinelli E, Simeone A, Acampora D, Gulisano M. Homeobox genes in the developing central nervous system. *Ann Genet* 1993;36(1):30-7

Bonne-Tamir B, DeStefano AL, Briggs CE, Adair R, Franklyn B, Weiss S, Korostishevsky M, Frydman M, Baldwin CT, Farrer LA. Linkage of congenital recessive deafness (gene DFNB10) to chromosome 21q22.3. *Am J Hum Genet*. 1996 Jun;58(6):1254-9.

Bradley A. Mining the mouse genome. *Nature*. 2002 Dec 5; 420(6915):512-4.

Bulfone A, Gattuso C, Marchitello A, Pardini C, Boncinelli E, Borsani G, Banfi S, Ballabio A. The embryonic expression pattern of 40 murine cDNAs homologous to Drosophila mutant genes (Dres): a comparative and topographic approach to predict gene function. *Hum Mol Genet*. 1998 Dec; 7 (13): 1997-2006.

Busciglio J, Pelsman A, Wong C, Pigino G, Yuan M, Mori H, Yankner BA. Altered metabolism of the amyloid beta precursor protein is associated with mitochondrial dysfunction in Down's syndrome. *Neuron*. 2002 Feb 28;33(5):677-88.

- Camacho Vanegas O, Bertini E, Zhang RZ, Petrini S, Minosse C, Sabatelli P, Giusti B, Chu ML, Pepe G. Ullrich scleroatonic muscular dystrophy is caused by recessive mutations in collagen type VI. *Proc Natl Acad Sci U S A*. 2001 Jun 19;98(13):7516-21.
- Carson JP, Thaller C, Eichele G. A transcriptome atlas of the mouse brain at cellular resolution. *Curr Opin Neurobiol*. 2002 Oct;12(5):562-5.
- Chaib H, Kaplan J, Gerber S, Vincent C, Ayadi H, Slim R, Munnich A, Weissenbach J, Petit C. A newly identified locus for Usher syndrome type I, USH1E, maps to chromosome 21q21. *Hum Mol Genet*. 1997 Jan;6(1):27-31.
- Dermitzakis ET, Reymond A, Lyle R, Scamuffa N, Ucla C, Deutsch S, Stevenson BJ, Flegel V, Bucher P, Jongeneel CV, Antonarakis SE. Numerous potentially functional but non-genic conserved sequences on human chromosome 21. *Nature*. 2002 Dec 5;420(6915):578-82.
- DeSilva U, Elnitski L, Idol JR, Doyle JL, Gan W, Thomas JW, Schwartz S, Dietrich NL, Beckstrom-Sternberg SM, McDowell JC, Blakesley RW, Bouffard GG, Thomas PJ, Touchman JW, Miller W, Green ED. Generation and comparative analysis of approximately 3.3 Mb of mouse genomic sequence orthologous to the region of human chromosome 7q11.23 implicated in Williams syndrome. *Genome Res*. 2002 Jan;12(1):3-15.
- Discala C, Ninnin M, Achard F, Barillot E, Vaysseix G. DBcat: a catalog of biological databases. *Nucleic Acids Res*. 1999 Jan 1;27(1):10-1.
- Dunham I, Shimizu N, Roe BA, Chissoe S, Hunt AR, Collins JE, Bruskiewich R, Beare DM, Clamp M, Smink LJ, Ainscough R, Almeida JP, Babbage A, Bagguley C, Bailey J, Barlow K, Bates KN, Beasley O, Bird CP, Blakey S, Bridgeman AM, Buck D, Burgess J, Burrill WD, O'Brien KP. The DNA sequence of human chromosome 22. *Nature*. 1999 Dec 2;402(6761):489-95.
- Engidawork E, Lubec G. Molecular changes in Down syndrome brain. *J Neurochem*. 2003; 84:895-904.
- Epstein, CJ. in *The Metabolic and Molecular Bases of Inherited Disease* (eds. Scriver CR, Beaudet AL, Sly WS, Valle D.) 749-794 (McGraw Hill, New York, 1995).
- Estabrooks LL, Rao KW, Donahue RP, Aylsworth AS. Holoprosencephaly in an infant with a minute deletion of chromosome 21(q22.3). *Am J Med Genet*. 1990 Jul;36(3):306-9.
- Frederikse PH, Ren XO. Lens defects and age-related fiber cell degeneration in a mouse model of increased AbetaPP gene dosage in Down syndrome. *Am J Pathol*. 2002 Dec;161(6):1985-90.
- Gawantka V, Pollet N, Delius H, Vingron M, Pfister R, Nitsch R, Blumenstock C, Niehrs C. Gene expression screening in *Xenopus* identifies molecular pathways, predicts gene function and provides a global view of embryonic patterning. *Mech Dev*. 1998 Oct;77(2):95-141.
- Ghadimi BM, Schrock E, Walker RL, Wangsa D, Jauho A, Meltzer PS, Ried T. Specific chromosomal aberrations and amplification of the AIB1 nuclear receptor coactivator gene in pancreatic carcinomas. *Am J Pathol*. 1999 Feb;154(2):525-36.
- Guipponi M, Tapparel C, Jousson O, Scamuffa N, Mas C, Rossier C, Hutter P, Meda P, Lyle R, Reymond A, Antonarakis SE. The murine orthologue of the Golgi-localized TPTE protein provides clues to the evolutionary history of the human TPTE gene family. *Hum Genet*. 2001 Dec;109(6):569-75.
- Hasle H, Clemmensen IH, Mikkelsen M. Risks of leukaemia and solid tumours in individuals with Down's syndrome. *Lancet*. 2000 Jan 15;355(9199):165-9.

- Hattori M, Fujiyama A, Taylor TD, Watanabe H, Yada T, Park HS, Toyoda A, Ishii K, Totoki Y, Choi DK, Groner Y, Soeda E, Ohki M, Takagi T, Sakaki Y, Taudien S, Blechschmidt K, Polley A, Menzel U, Delabar J, Kumpf K, Lehmann R, Patterson D, Reichwald K, Rump A, Schillhabel M, Schudy A, Zimmermann W, Rosenthal A, Kudoh J, Schibuya K, Kawasaki K, Asakawa S, Shintani A, Sasaki T, Nagamine K, Mitsuyama S, Antonarakis SE, Minoshima S, Shimizu N, Nordsiek G, Hornischer K, Brant P, Scharfe M, Schon O, Desario A, Reichelt J, Kauer G, Blocker H, Ramser J, Beck A, Klages S, Hennig S, Riesselmann L, Dagand E, Haaf T, Wehrmeyer S, Borzym K, Gardiner K, Nizetic D, Francis F, Lehrach H, Reinhardt R, Yaspo ML. The DNA sequence of human chromosome 21. *Nature*. 2000 May 18;405(6784):311-9.
- Head E, Lott IT, Cribbs DH, Cotman CW, Rohn TT. Beta-amyloid deposition and neurofibrillary tangle association with caspase activation in Down syndrome. *Neurosci Lett*. 2002 Sep 13;330(1):99-103.
- Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. *Genome Res*. 1996 Oct;6(10):986-94.
- Herzig U, Cadenas C, Sieckmann F, Sierralta W, Thaller C, Visel A, Eichele G. Development of high-throughput tools to unravel the complexity of gene expression patterns in the mammalian brain. *Novartis Found Symp* 239, 129-46 (2001).
- Higuchi I, Shiraishi T, Hashiguchi T, Suehara M, Niiyama T, Nakagawa M, Arimura K, Maruyama I, Osame M. Frameshift mutation in the collagen VI gene causes Ullrich's disease. *Ann Neurol*. 2001 Aug;50(2):261-5.
- Ichikawa H, Hosoda F, Arai Y, Shimizu K, Ohira M, Ohki M. A NotI restriction map of the entire long arm of human chromosome 21. *Nat Genet*. 1993 Aug;4(4):361-6.
- Jobsis GJ, Keizers H, Vreijling JP, de Visser M, Speer MC, Wolterman RA, Baas F, Bolhuis PA. Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. *Nat Genet*. 1996 Sep;14(1):113-5.
- Kohno T, Kawanishi M, Matsuda S, Ichikawa H, Takada M, Ohki M, Yamamoto T, Yokota J. Homozygous deletion and frequent allelic loss of the 21q11.1-q21.1 region including the ANA gene in human lung carcinoma. *Genes Chromosomes Cancer*. 1998 Mar;21(3):236-43.
- Kola I, Hertzog PJ. Animal models in the study of the biological function of genes on human chromosome 21 and their role in the pathophysiology of Down syndrome. *Hum Mol Genet*. 1997;6(10):1713-27.
- Lander ES and the Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001 Feb 15;409(6822):860-921.
- Lercher MJ, Urrutia AO, Hurst LD. Clustering of housekeeping genes provides a unified model of gene order in the human genome. *Nat Genet*. 2002 Jun;31(2):180-3.
- Li Q, Harju S, Peterson KR. Locus control regions: coming of age at a decade plus. *Trends Genet*. 1999 Oct;15(10):403-8.
- Linggi B, Muller-Tidow C, van de Locht L, Hu M, Nip J, Serve H, Berdel WE, van der Reijden B, Quelle DE, Rowley JD, Cleveland J, Jansen JH, Pandolfi PP, Hiebert SW. The t(8;21) fusion protein, AML1 ETO, specifically represses the transcription of the p14(ARF) tumor suppressor in acute myeloid leukemia. *Nat Med*. 2002 Jul;8(7):743-50.
- McInnis MG, Chakravarti A, Blaschak J, Petersen MB, Sharma V, Avramopoulos D, Blouin JL, Konig U, Brahe C, Matise TC. A linkage map of human chromosome 21:43 PCR markers at average intervals of 2.5cM. *Genomics*. 1993 Jun;16(3):562-71.

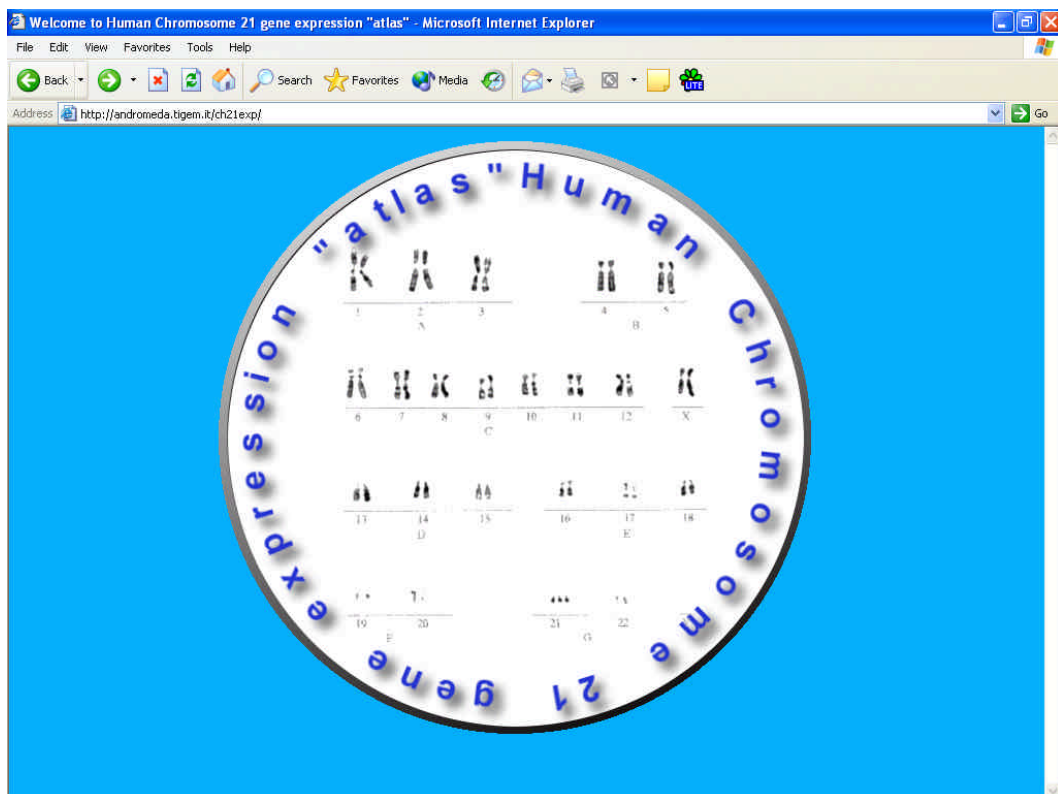
- Metzler M, Legendre-Guillemain V, Gan L, Chopra V, Kwok A, McPherson PS, Hayden MR. HIP1 functions in clathrin-mediated endocytosis through binding to clathrin and adaptor protein 2. *J Biol Chem*. 2001 Oct 19;276(42):39271-6.
- Motonaga K, Itoh M, Becker LE, Goto Y, Takashima S. Elevated expression of beta-site amyloid precursor protein cleaving enzyme 2 in brains of patients with Down syndrome. *Neurosci Lett*. 2002 Jun 21;326(1):64-6.
- Neidhardt L, Gasca S, Wertz K, Obermayr F, Worpenberg S, Lehrach H, Herrmann BG. Large-scale screen for genes controlling mammalian embryogenesis, using high-throughput gene expression analysis in mouse embryos. *Mech Dev*. 2000 Nov;98(1-2):77-94.
- Ohgaki K, Iida A, Kasumi F, Sakamoto G, Akimoto M, Nakamura Y, Emi M. Mapping of a new target region of allelic loss to a 6-cM interval at 21q21 in primary breast cancers. *Genes Chromosomes Cancer*. 1998 Nov;23(3):244-7.
- Pennacchio LA, Rubin EM. Genomic strategies to identify mammalian regulatory sequences. *Nat Rev Genet*. 2001 Feb;2(2):100-9.
- Pletcher MT, Wiltshire T, Cabin DE, Villanueva M, Reeves RH. Use of comparative physical and sequence mapping to annotate mouse chromosome 16 and human chromosome 21. *Genomics*. 2001 May 15;74(1):45-54.
- Pucharcos C, Casas C, Nadal M, Estivill X, de la Luna S. The human intersectin genes and their spliced variants are differentially expressed. *Biochim Biophys Acta*. 2001 Oct 31;1521(1-3):1-11.
- Reeves RH. Functional genomics: a time and place for every gene. *Nature*. 2002 Dec 5;420(6915):518-9.
- Reeves RH, Irving NG, Moran TH, Wohn A, Kitt C, Sisodia SS, Schmidt C, Bronson RT, Davisson MT. A mouse model for Down syndrome exhibits learning and behaviour deficits. *Nat Genet*. 1995 Oct;11(2):177-84.
- Reeves RH, Yao J, Crowley MR, Buck S, Zhang X, Yarowsky P, Gearhart JD, Hilt DC. Astrocytosis and axonal proliferation in the hippocampus of S100b transgenic mice. *Proc Natl Acad Sci U S A* 1994 Jun 7;91(12):5359-63
- Reymond A, Friedli M, Henrichsen CN, Chapot F, Deutsch S, Ucla C, Rossier C, Lyle R, Guipponi M, Antonarakis SE. From PREDs and open reading frames to cDNA isolation: revisiting the human chromosome 21 transcription map. *Genomics*. 2001 Nov;78(1-2):46-54.
- Reymond A, Camargo AA, Deutsch S, Stevenson BJ, Parmigiani RB, Ucla C, Bettoni F, Rossier C, Lyle R, Guipponi M, de Souza S, Iseli C, Jongeneel CV, Bucher P, Simpson AJ, Antonarakis SE. Nineteen additional unpredicted transcripts from human chromosome 21. *Genomics*. 2002a Jun;79(6):824-32.
- Reymond A, Marigo V, Yaylaoglu MB, Leoni A, Ucla C, Scamuffa N, Caccioppoli C, Dermitzakis ET, Lyle R, Banfi S, Eichele G, Antonarakis SE, Ballabio A. Human chromosome 21 gene expression atlas in the mouse. *Nature*. 2002b Dec 5;420(6915):582-6.
- Ringwald M, Eppig JT, Richardson JE. GXD: integrated access to gene expression data for the laboratory mouse. *Trends Genet*. 2000 Apr; 16(4):188-90.
- Sago H, Carlson EJ, Smith DJ, Kilbridge J, Rubin EM, Mobley WC, Epstein CJ, Huang TT. Ts1Cje, a partial trisomy 16 mouse model for Down syndrome, exhibits learning and behavioral abnormalities. *Proc Natl Acad Sci U S A*. 1998 May 95:6256-6261.

- Sakata K, Tamura G, Nishizuka S, Maesawa C, Suzuki Y, Iwaya T, Terashima M, Saito K, Satodate R. Commonly deleted regions on the long arm of chromosome 21 in differentiated adenocarcinoma of the stomach. *Genes Chromosomes Cancer*. 1997 Apr;18(4):318-21.
- Satge D, Sasco AJ, Carlsen NL, Stiller CA, Rubie H, Hero B, de Bernardi B, de Kraker J, Coze C, Kogner P, Langmark F, Hakvoort-Cammel FG, Beck D, von der Weid N, Parkes S, Hartmann O, Lippens RJ, Kamps WA, Sommelet D. A lack of neuroblastoma in Down syndrome: a study from 11 European countries. *Cancer Res*. 1998a Feb 1;58(3):448-52.
- Satge D, Sommelet D, Geneix A, Nishi M, Malet P, Vekemans M. A tumor profile in Down syndrome. *Am J Med Genet*. 1998b Jul 7;78(3):207-16.
- Schreiber SL, Bernstein BE. Signaling network model of chromatin. *Cell*. 2002 Dec 13;111(6):771-8.
- Sertie AL, Sossi V, Camargo AA, Zatz M, Brahe C, Passos-Bueno MR. Collagen XVIII, containing an endogenous inhibitor of angiogenesis and tumor growth, plays a critical role in the maintenance of retinal structure and in neural tube closure (Knobloch syndrome). *Hum Mol Genet*. 2000 Aug 12;9(13):2051-8.
- Sertie AL, Quimby M, Moreira ES, Murray J, Zatz M, Antonarakis SE, Passos-Bueno MR. A gene which causes severe ocular alterations and occipital encephalocele (Knobloch syndrome) is mapped to 21q22.3. *Hum Mol Genet*. 1996 Jun;5(6):843-7.
- Smithies O. Animal models of human genetic diseases. *Trends Genet*. 1993 Apr;9(4):112-6.
- Speel EJ, Saremaslani P, Roth J, Hopman AH, Komminoth P. Improved mRNA in situ hybridization on formaldehyde-fixed and paraffin-embedded tissue using signal amplification with different haptenized tyramides. *Histochem Cell Biol*. 1998 Dec; 110 (6): 571-7.
- Stathopoulos A, Van Drenth M, Erives A, Markstein M, Levine M. Whole-genome analysis of dorsal-ventral patterning in the *Drosophila* embryo. *Cell*. 2002 Nov 27;111(5):687-701.
- Strachan T, Abitbol M, Davidson D, Beckmann JS. A new dimension for the human genome project: towards comprehensive expression maps. *Nat Genet*. 1997 Jun;16(2):126-32.
- Sumarsono SH, Wilson TJ, Tymms MJ, Venter DJ, Corrick CM, Kola R, Lahoud MH, Papas TS, Seth A, Kola I. Down's syndrome-like skeletal abnormalities in *Ets2* transgenic mice. *Nature*. 1996 Feb 8; 379(6565): 534-7.
- van Belle H. Kinetics and inhibition of alkaline phosphatases from canine tissues. *Biochim Biophys Acta*. 1972 Nov 10;289(1):158-68.
- Visel A, Ahdidan J and Eichele G. A gene expression map of the mouse brain: Genepaint.org - a database of gene expression patterns. In: *Neuroscience Databases: A Practical Guide*. Koetter R., ed. (Kluwer Academic Publishers, Boston, Dordrecht, London), 2002, 19-36.
- Wiltshire T, Pletcher M, Cole SE, Villanueva M, Birren B, Lehoczky J, Dewar K, Reeves RH. Perfect conserved linkage across the entire mouse chromosome 10 region homologous to human chromosome 21. *Genome Res*. 1999 Dec;9(12):1214-22.
- Yamamoto N, Uzawa K, Miya T, Watanabe T, Yokoe H, Shibahara T, Noma H, Tanzawa H. Frequent allelic loss/imbalance on the long arm of chromosome 21 in oral cancer: evidence for three discrete tumor suppressor gene loci. *Oncol Rep*. 1999 Nov-Dec;6(6):1223-7.
- Zreiqat H, Sungaran R, Howlett CR, Markovic B. Quantitative aspects of an in situ hybridization procedure for detecting mRNAs in cells using 96-well microplates. *Mol Biotechnol*. 1998 Oct; 10 (2): 107-13.

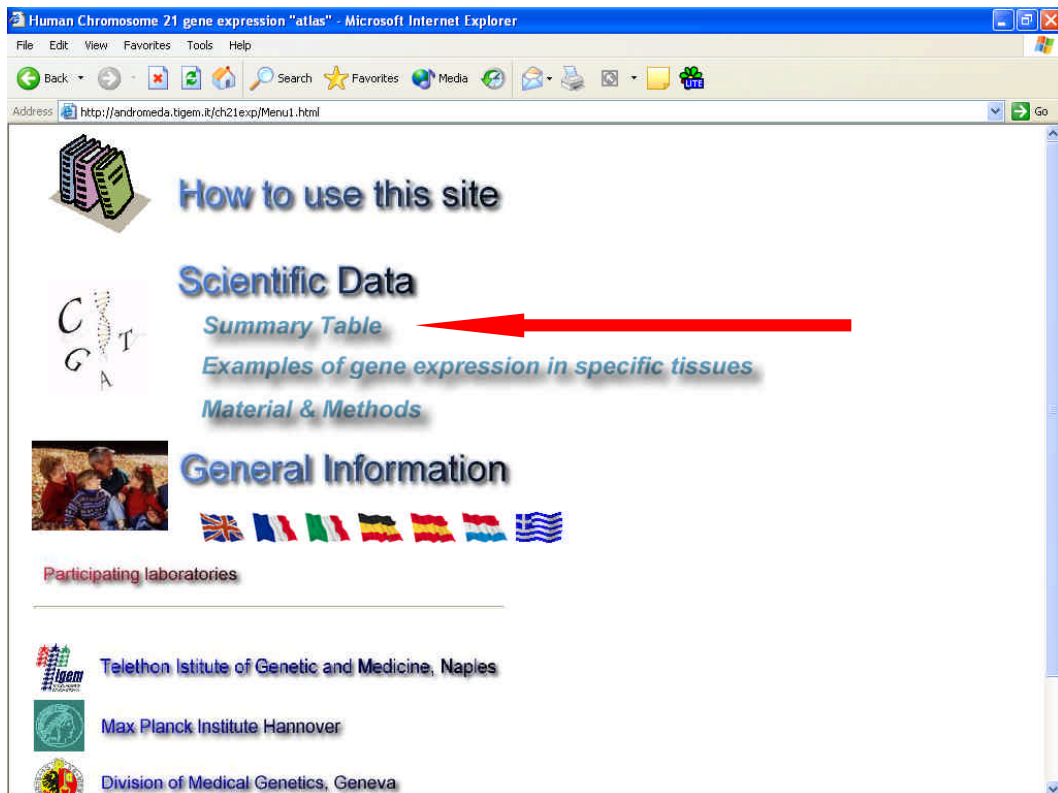
VII. Appendix

7.1 Brief introduction to www.tigem.it/ch21exp/

- After entering the world wide web address www.tigem.it/ch21exp/ the entrance page to the “human chromosome 21 gene expression atlas” will open up on the web browser



- After you enter the atlas, you can access the scientific data by clicking onto the various headings under “scientific data”,
 - summary table
 - example of gene expression in specific tissues
 - materials and methodsA guide on how to use this site can be accessed under the heading “how to use this site”.



- After clicking on the “summary table” link the page showing the summary table which has all the HC21 genes listed in chromosomal order and the accession numbers for their murine orthologues opens. When one of the HC21 genes are clicked upon, the expression data belonging to the corresponding murine orthologue is accessible. For example click on the link to the *BTG3* gene.

HUMAN CHROMOSOME 21 GENES AND MURINE ORTHOLOGUES

Number	Gene name and aliases	Hs Accession ¹	Mm Accession ¹	Status in mouse ²
1	TPTE/PTEN2	NM_013315	AJ311311	done Expression Atlas
2	C21orf15	AY040090		no murine orthologue
3	PRED5	AP001660		no murine probes
4	REM11	AP001660	BB633189	done Expression Atlas
5	PRED6	AP001660		no murine probes
6	STCH	NM_006948	AK021006	done Expression Atlas
7	SAMSN1	NM_022136	NM_023380	done Expression Atlas
8	NRIP1/RIP140	NM_003489	NM_008735	done Expression Atlas
9	USP25	NM_013396	NM_013918	done Expression Atlas
10	C21orf34/C21orf35	AP001666		no murine probes
11	CXADR	NM_001338	NM_009988	done Expression Atlas
12	BTG3			done Expression Atlas

Genbank accession(s) number(s)
 2. orthologous gene status in mouse:
 done Expression Atlas: ■
 done RT-PCR panel only: ■
 no murine probes: ■
 no murine orthologue: ■

7.2 Brief introduction to www.genepaint.org

- After entering the world wide web address www.genepaint.org into your browser the home page to the GenePaint database will be produced. The gene symbol for the gene of interest is entered and then the “go” button is pressed to start the query (e.g. *Sim2*). Then the next screen will pop up showing the stages that have been studied for this gene. The database contains embryonic day 14.5 whole embryo sections, and often also embryonic day 15.5 head sections, P7 and adult mice brain sections.

Enter gene name and click go to initiate query

Unbenanntes Dokument - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media

Address <http://www.genepaint.org/frameset.html> Go

gene paint.org

GenePaint Home Quick guide to GenePaint Brain Maps Advanced Search

Search GenePaint

Sim2 go

enter gene name, accession number or GenePaint set ID

Advanced search

GenePaint Home

Manual of GenePaint

Evaluation of Expression

RNA in situ Hybridization

Zoom Viewer

QUICK GUIDE

Brain Maps

Request

Funding of GenePaint

User Name

Password go

Gene Directory

30.Apr.2003

New Genes

Mf2

cryptochrome 1_photolyase-like

choline acetyltransferase

drcadian locomoter output cycles kaput

choline acetyltransferase

junction cell adhesion molecule 2

Iroquois related homeobox 1

WDR9

NRP1

Synj1

Bdnf

dopamine receptor D1A

Notch gene homolog 3_Drosophila

LIM homeobox protein 6

melanocortin 4 receptor

GenePaint.org depicts gene expression patterns in the mouse. Patterns were determined by RNA in situ hybridization on tissue sections. The goal of GenePaint.org is to serve as an atlas of expression patterns of all genes.

Expression data are retrieved through gene names, sequence accession number, a BLAST search or a GenePaint set identification number. In the near future, site and intensity of gene expression will be annotated, and hence expression data will also be searchable using anatomical features.

For viewing expression patterns, GenePaint.org provides a "virtual microscope" tool that enables zooming into images.

GenePaint.org users can suggest genes to be included into the database using a [request form](#).

The home page of GenePaint.org features a password-protected section that holds expression data undergoing quality control and hence not yet released for general viewing.

The development of the technology behind GenePaint was initiated by Dr. Gregor Eichele at the Max-Planck Institute in Hannover, Germany. Data collection was begun in mid

MAX-PLANCK-INSTITUT HANNOVER

Please click here for download of the required plugin
Plugin for Windows
Plugin for Mac OS

javascript:startquery(true,false);

Internet

- The results window allows an overview of stages that have been examined for the query gene. Certain advanced query options are also presented in this window and allow e.g. entry of DNA sequence which will then be blasted against the sequences of the probes used for ISHs within the database. The blast results are presented from the highest to the lowest score. The options boxes at the bottom left corner allow the queries to be limited to certain developmental stages.

The screenshot shows the GenePaint.org website interface. The search bar contains the text "Sim2". The results table shows one hit for Set ID MY82, Accession no. NM_011377, Gene Name Sim2, Tissue Embryo, Stage E14.5, Strain NMRI. Callouts provide additional context: "Instead of a gene name a sequence may be used for the querying of the database" points to the search input; "The setviewer initiates the data viewing window" points to the "Set Viewer" link; "The query could be limited to a certain developmental stage" points to the "Stages" dropdown menu.

Query

Search for a gene (probe, gene symbols, gene names, GenePaint Set ID's. Names with spaces are entered between quotation marks ("Pax6").

Sim2

Go

Clear

Reset

BLAST with Sequence below

Identity

high(>200)

medium(>100)

low(>50)

minimal score

Additional Restrictions

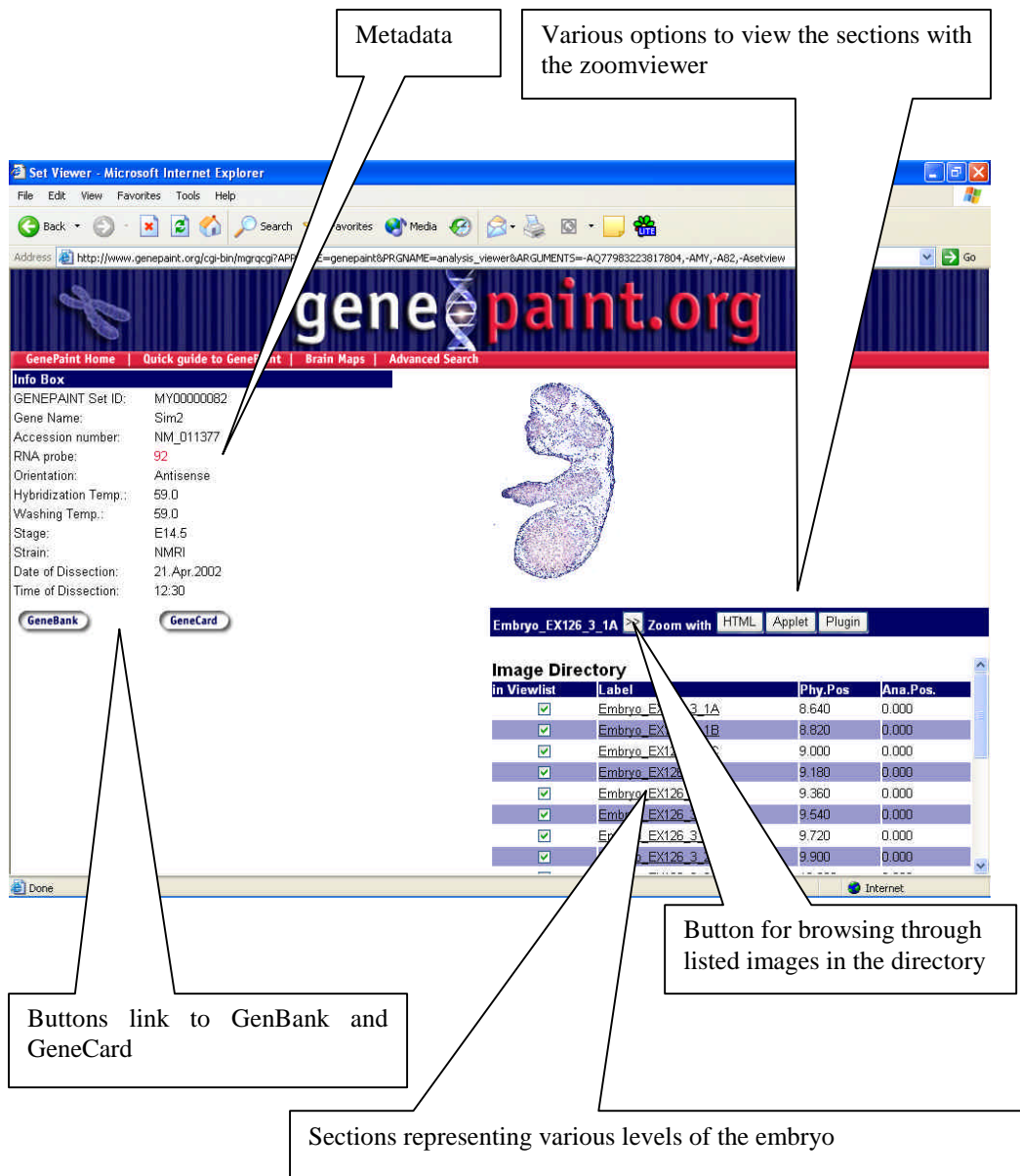
Stages	Strains	Tissue
any	any	any
P56	C57BL/6	Brain
P7	NMRI	Embryo
E15.5	SV/B16/29	Head

Results

Set ID	Accession no.	Gene Name	Tissue	Stage	Strain	Views
MY82	NM_011377	Sim2	Embryo	E14.5	NMRI	Thumbnail Set Viewer

Show 20 1 Hits Page 1 of 1 select page: 1

- Results are viewed by clicking the setviewer option. The resulting window shows an interface with metadata (left side), presenting experimental conditions and sequence information. Below the infobar are two buttons that link to the GenBank and GeneCard databases. On the right side of the screen there is a thumbnail of the section of interest, and by choosing one of the viewer options (html format, applet interface or a downloadable plug-in software) one is able to examine the section shown in the thumbnail with the help of an “electronic microscope”. The images can rapidly be browsed using the forward and reverse arrow button.



- Example depicting a section that is being viewed with the applet viewer. The small window that has popped up has various tools at the bottom that assist in viewing. The (+) button enables zooming in, the (-) button enables zooming out, the hand allows panning, and the magnification button (+) brings up a small window allowing a magnification of a specific section within the total image.

The screenshot shows a Microsoft Internet Explorer window displaying the GenePaint applet. The main window shows a histological section of a mouse embryo. Below the image is a toolbar with icons for zooming in (+), zooming out (-), panning (hand), and a magnification window (+). Below the toolbar, the text 'GenePaint Set ID: MY82 Gene: Sim2 Stage: E14.5' is visible. At the bottom of the window, a table lists gene expression data for various embryonic regions.

<input checked="" type="checkbox"/>	Embryo_EX126_3_5D	11.340	0.000
<input checked="" type="checkbox"/>	Embryo_EX126_3_5A	11.520	0.000
<input checked="" type="checkbox"/>	Embryo_EX126_3_5B	11.700	0.000
<input checked="" type="checkbox"/>	Embryo_EX126_3_5C	11.880	0.000
<input checked="" type="checkbox"/>	Embryo_EX126_3_5D	12.060	0.000

Tools that assist in viewing
 (+) zoom in
 (-) zoom out
 (the hand) panning
 magnification window

Curriculum vitae

Personal Information

Name: Murat Burak YAYLAOGLU
Name of Father: Ahmet
Name of Mother: Nezahat
Date and Place of Birth: July 7, 1971 - Ankara, Turkey
Nationality: Turkish

Education

Doctoral Research:		Max-Planck-Institute. Dept. of Molecular Embryology, Hannover, Germany.
Master of Science Degree in Biology:	1994 to 1997	Middle East Technical University. Ankara, Turkey. Department of Biological Sciences. (CGPA 3.25) (Date of Graduation: 31 January 1997) Theme of research: Development of a bone substitute, its use in drug release and its <i>in vitro</i> evaluation with cultured chondrocytes.
Bachelor of Science Degree in Biology:	1989 to 1993	Hacettepe University. Ankara, Turkey. Faculty of Science. Department of Biology. (CGPA 3.02) (Date of graduation: 23 July 1993)
High School:	1983 to 1988	Ankara Atatürk Anadolu Secondary High School, Ankara, Turkey.
Primary School:	1976 to 1982	Auburn North Primary School, Sydney, Australia

Conferences

- Yaylaoglu MB, Korkusuz P, Ors U, et al. The development of a novel composite bone substitute for use in drug release 4th International Symposium on Biomedical Science and Technology (Biomed 4), SEP 15-17, 1997 Biomedical science and technology recent developments in the pharmaceutical and medical sciences , 205-215, 1998
- 13th Biennial Meeting of the International Society for Developmental Neuroscience. 28.July.2000.
- Oral presentation; "Systematic Gene Expression Pattern Analysis on a Large Scale Using Automated High-throughput *In Situ* Hybridization" German Human Genome Meeting 2001 (November 07 - 09, 2001), Braunschweig, Germany.
- Poster presentation titled; " Visualizing Genome-wide Gene Expression: from High-throughput Tools (GenePaint) to the Internet (GenePaint.org)" The first joint Meeting of the NGFN and the DHGP, (November 17 - 19, 2002) Berlin, Germany.

Scholarships

- 1995 Three month DAAD (German Academic Exchange Service) Scholarship.
- 2003 NIH grant award for the C6 meeting entitled "Functional Genomics: Global Analysis of Complex Biological Systems".(Department of Health and Human Services, National Institutes of Health, National Institute of General Medical Studies: Grant Number 1 R13 GM067252-01).

Publications

- **Yaylaoglu, MB et al.** "Development of a calcium phosphate-gelatin composite as a bone substitute and its use in drug release". *Biomaterials*. 1999 Apr; 20 (8): 711 - 719. (M.Sc. Degree)
- **Yaylaoglu, MB et al.** "A novel osteochondral implant". *Biomaterials*. 1999 Aug; 20, (16): 1513-20. (M.Sc. Degree)
- Kraszucka K, Burfeind P, Nayernia K, Kohler M, Schmid M, **Yaylaoglu M**, Engel W. "Developmental stage- and germ cell-regulated expression of a calcium-binding protein mRNA in mouse Sertoli cells". *Mol Reprod Dev*. 1999 Nov; 54 (3): 232-243. (Predoctorate research)
- Moeller C, **Yaylaoglu MB**, Alvarez-Bolado G, Thaller C, Eichele G. "Murine *Lix1*, a novel marker for *substantia nigra*, cortical layer 5, and hindbrain structures". *Gene Expression Patterns* 2002 October; 1, 199-203. (Ph.D. Degree)
- **Reymond A⁺, Marigo V⁺, Yaylaoglu MB⁺**, Leoni A, UCLA C, Scamuffa N, Caccioppoli C, Dermitzakis ET, Lyle R, Banfi S, Eichele G, Antonarakis SE, Ballabio. "A. Human chromosome 21 gene expression atlas in the mouse". *Nature* 2002 Dec; 420, 582-586 (Ph.D. Degree) + **shared first authorship.**
- **Marigo V⁺, Yaylaoglu MB⁺, UCLA C⁺**, Leoni A, Merla G, Scamuffa N, Titmus A, Ballabio A, Thaller C, Antonarakis SE, Eichele G, Reymond A. "The Atlas of Expression of the Williams-Beuren Contiguous Gene Disorder". *In preparation* (Ph.D. Degree) + **shared first authorship.**