

On HIRA, Chromosome 22q11 and CATCH22

Stellingen behorende bij het proefschrift

On HIRA, Chromosome 22q11 and CATCH22

Laurens Wilming, Hinxton, November 1999

The thesis

- 1: The large size of deletions in CATCH22 patients is not necessarily an argument for it being a multigene syndrome, as the underlying mutational mechanism could be responsible for the size.
This thesis.
- 2: Many obstacles in the areas of expertise, software and CPU power and time are encountered by small research groups involved in large scale genome analysis. Therefore, analysis is best left to large genome centers.
This thesis.
- 3: DNA based diagnostics are essential for supporting clinical observations with regards to DiGeorge craniofacial abnormalities.
This thesis.
- 4: Even though we are close to knowing the majority of the human genes, the challenge of determining the function of their products and relevance to diseases and syndromes remains as daunting as ever.
This thesis.

The brain

- 5: Gödel's incompleteness theory can be applied to the consciousness of our inner self and the awareness of our thought processes.
K Gödel (1931) I. Monatshefte für Math. u. Physik 38: 173-198; K Gödel (1992): On Formally Undecidable Propositions of Principia Mathematica and Related Systems. Dover, New York, NY; DR Hofstadter (1989): Gödel, Escher, Bach: An Eternal Golden Braid. Vintage Books, New York, NY.
- 6: The unconscious homunculus: an oxymoron?
F Crick and C Koch (1999): The Unconscious Homunculus. In: T Metzinger (ed) (1999): The Neural Correlates of Consciousness. MIT press, Cambridge, MA.
- 7: "An evolutionary consciousness is a unique means by which energy, using stable physico-chemical and biological codal processes, develops a means of increasing the pragmatics of both laws of thermodynamics by the development of hierarchical levels of metareferential code systems".
E Taborsky (1999) Biosystems 51: 153-168.
- 8: Man is a victim of its brain.
Victim of the Brain. Film by Piet Hoenderdos after DR Hofstadter and DC Dennett (1982): The Mind's I. Penguin.

The computer

- 9: wysiwyg is most often actually wysiwywywag (what you see is what you wish you would actually get), though sometimes it is wygiwywyws (what you get is what you wish you would see).

The car and the computer

- 10: A small sampling shows that a disproportionate number of Mac users are Saab drivers with creative jobs. They adhere to the motto: 'Think Different. Pay More.'
- 11: Compared to the MacOS, Windows has the subtlety and ergonomics of a Filipino Taxi compared to an automobile from the Volkswagen-Audi stable.

The rest

- 12: The dangers of isolation and group behavior (forced or self-imposed) is witnessed by the high incidence of certain hereditary diseases in isolated groups like Afrikaanders, Ashkenazi and Cephardic Jews, Fins and Zeeuwen.
AD Auerbach (1997), Genet Test 1: 27-33; J Rosendorff et al. (1987), Am J Med Genet 27: 793-797; L Peltonen et al. (1999), Hum Mol Genet 8: 1913-1923.
- 13: Looking at embryonic development in 3D adds a whole new dimension to our knowledge.
J Scarborough et al. (1997), J Anat 191: 117-122.
- 14: Life is a fatal condition we inherit from our parents.
- 15: Thinking that honey made of pollen from genetically modified plants is harmful is like thinking that giving a blind man an anarchist manual is more harmful than giving him an innocent children's book.

On HIRA, Chromosome 22q11 and CATCH22

Over HIRA, Chromosoom 22q11 en CATCH22

Proefschrift

ter verkrijging van de graad van doctor aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. dr. P.W.C. Akkermans M.A.
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 19 januari 2000 te 11:45 uur

door

LAURENS GERARDUS WILMING

geboren te Utrecht

Promotiecommissie

Promotores: Prof. dr. F.G. Grosveld
Prof. dr. D. Tibboel

Copromotor: Dr. J.H.C. Meijers

Overige leden: Prof. dr. E.A. Dzierzak
Prof. dr. B.A. Oostra
Prof. dr. R. Zeller

Het in dit proefschrift beschreven onderzoek is bewerkt binnen de vakgroep Celbiologie en Genetica, Faculteit der Geneeskunde en Gezondheidswetenschappen, Erasmus Universiteit Rotterdam in samenwerking met het Instituut Kinderheilkunde van het Sophia Kinderziekenhuis. Financiële ondersteuning van het onderzoek werd ondermeer verzorgd door de Sophia Stichting voor Wetenschappelijk Onderzoek, de Nederlandse Hartstichting en de Europese Gemeenschap.

Voor mijn vader
Ter nagedachtenis aan mijn moeder



Contents

page section

iii	0 CONTENTS
1	I..IV INTRODUCTION
3	I INTRODUCING THE INTRODUCTION
3	I-1 GENERAL
4	I-2 SCOPE OF THIS THESIS
4	II INTRODUCTION TO THE NEURAL CREST AND THE PHARYNGEAL ARCHES
4	II-1 GENERAL
6	II-2 THE RHOMBENCEPHALON IS SEGMENTED INTO ODD AND EVEN NUMBERED RHOMBOMERES WITH SEPARATE IDENTITIES
8	II-3 FATES, DESTINATIONS AND CONTRIBUTIONS OF NEURAL CREST
8	II-3-1 Neural Crest Cells Are Multipotent: Summary of In Vitro Results
9	II-3-2 Fate Mapping Shows the Myriad Contributions of the Neural Crest
10	II-3-3 Cranial Neural Crest and the First Two Branchial Arches Produce Most of the Face
12	II-3-4 Thymus, Thyroid and Parathyroids Depend on Neural Crest for Proper Development
13	II-3-5 Remodeling of the Pharyngeal Arch Artery System Is Crucial for Anterior Cardiovascular Development
15	II-4 MOLECULAR GENETICS OF NEURAL CREST
15	II-4-1 A Complex Cascading Network of Genes Is Involved in Neural Crest Patterning, Development and Fate Determination
16	II-4-1-1 Homeobox Genes
17	II-4-1-2 Endothelin System
18	II-4-2 Neural Crest Cell Migration Is Regulated by Various ECM Components, but also by Soluble Factors and their Receptors
20	II-5 DEPENDING ON THE AXIAL LEVEL, EXTIRPATED NEURAL CREST IS COMPENSATED FOR BY REMAINING NEURAL TISSUE
21	III DISEASES AND SYNDROMES
21	III-1 SYNDROMES RELATED TO HINDBRAIN NEURAL CREST MALDEVELOPMENT
22	III-2 SINGLE GENE AND CONTIGUOUS GENE HAPLO-INSUFFICIENCY OR AUTOSOMAL SEGMENTAL ANEUPLOIDY SYNDROMES
23	III-3 CHROMOSOME 22Q11 DELETION SYNDROME: NEURAL CREST AND CHROMOSOME 22
24	III-3-1 Various Disorders Are Associated with Chromosome 22 Rearrangements
24	III-3-2 Chromosome 22q11 Anomalies Are Associated with a Variety of Cardiovascular and Craniofacial Abnormalities, Many of Which Are Cranial Neural Crest Related
25	III-3-2-1 Cardiac Defects
26	III-3-2-2 Abnormal Facies
26	III-3-2-3 Thymic Hypoplasia or Aplasia
26	III-3-2-4 Cleft Palate
27	III-3-2-5 Hypocalcemia
27	III-3-2-6 22q11 Aberrations
28	III-3-3 A Clear Link Between Chromosome 22q11 and Neural Crest Development
28	III-3-4 Syndrome, Developmental Field Defect, Anomaly, Malformation
29	IV CONCLUDING THE INTRODUCTION
31	V..VIII RESULTS & DISCUSSION
33	V INTRODUCTION TO THE EXPERIMENTAL WORK
35	VI LOW RESOLUTION MAP: CONSTRUCTION OF FOURCOSMID CONTIGS ON 22Q11
35	VI-1 ANONYMOUS SINGLE COPY PROBE NB84 LOCATES TO 22Q11
36	VI-2 GRIDDED CHROMOSOME 22 COSMID LIBRARY

page	section
37	VI-3 NB84 RECOGNIZES FOUR COSMIDS FROM TWO SITES ON CHROMOSOME 22
37	VI-4 M51 IS LOCATED CENTROMERIC OF M56 BUT BOTH ARE LOCATED WITHIN THE DGCR
38	VI-5 M51 COSMID CONTIG
38	VI-6 M56 COSMID CONTIG
39	VI-7 SINGLE COPY PROBES CONFIRM THE ARRANGEMENTS OF THE CONSENSUS CONTIGS
40	VI-8 HIRA/TUPLE1 AND HD7K CONTIGS
40	VI-9 YACS
40	VI-10 cDNA LIBRARY SCREENINGS WITH CONTIG COSMIDS: M51 AND M56 CONTIG COSMIDS DO NOT CONTAIN ABUNDANT OR COMMONLY EXPRESSED SEQUENCES
41	VI-11 DISCUSSION
41	VII HIGH RESOLUTION MAP: MAPPING OF NEW GENES IN THE 22Q11 REGION
42	VII-1 22q11 CONTIG 1
42	VII-1-1 Established Transcription Units
43	VII-1-2 Notable Features
44	VII-2 22q11 CONTIG 2
46	VII-3 22q11 CONTIG 3
46	VII-3-1 Established Transcription Units
46	VII-3-2 New Transcription units
48	VII-4 22q11 CONTIG 4
48	VII-4-1 Established Transcription Units
48	VII-4-2 New Transcription units
52	VII-5 22q11 CONTIG 5
52	VII-5-1 Established Transcription Units
52	VII-5-2 New Transcription units
53	VII-5-3 Notable Features
53	VII-6 MINOR INTERGENIC AND INTRONIC TRANSCRIPTS
53	VII-7 LOW COPY REPEATS
53	VII-8 GENERAL DISCUSSION OF DETAILED ANALYSIS OF DGCR GENOMIC SEQUENCE
56	VIII ANALYSIS OF HIRA (DGCR1, TUPLE1), A CATCH22 CANDIDATE GENE
56	VIII-1 HUMAN HIRA
57	VIII-2 MURINE HIRA cDNA CLONES CODE FOR A PROTEIN WITH HOMOLOGY TO HUMAN CHROMATIN ASSEMBLY FACTOR 1A AND MULTIPLE ALIGNMENT SUGGESTS THE PRESENCE OF THREE FUNCTIONAL REGIONS IN HIRA
60	VIII-3 TWO HIRA TRANSCRIPTS ARE PRESENT FROM GASTRULATION STAGE ONWARD
61	VIII-4 HIRA IS UBIQUITOUSLY EXPRESSED, WITH ENHANCED EXPRESSION IN PHARYNGEAL ARCHES, NEURAL FOLDS, CIRCUMPHARYNGEAL NEURAL CREST AND LIMBS
63	VIII-5 cDNA CLONES 12B AND 12E REPRESENT SPLICE VARIANTS WITH EXONS 3A AND 11A
67	IX DISCUSSION & CONCLUSIONS
69	IX-1 INSIGHT
72	IX-2 OUTLOOK
72	IX-2-1 Bioinformatics and Genome Projects Will Be a Major Boon to Medical Genetic Research
73	IX-2-2 Reverse Genetics: Be Aware of the Pitfalls of Using the Mouse as a Model Organism
74	IX-2-3 Studying Other Congenital Syndromes Provides Insight into Complex Genotype-Phenotype Relationships
75	IX-2-4 Ongoing Research into HIRA Reveals that It Is Almost Certainly Involved in Chromatin Interactions

page	section
76	IX-2-5 <i>Is the 'DiGeorge Gene' Still to be Identified?</i>
76	IX-2-6 <i>Studying CATCH22 Phenocopies and Limb Development May Yield Valuable Clues About the Regulatory Pathways in Craniofacial Development</i>
77	IX-2-7 <i>Clinical Contribution To Our Knowledge Is Expanding And Improving</i>
79	X MATERIALS & METHODS
81	X-1 PROBES
81	X-2 PATIENTS AND CELL LINES
81	X-3 KARYOTYPING AND FISH
82	X-4 SOUTHERN ANALYSIS
82	X-4-1 Human DNA
82	X-4-2 Zoo Blot
82	X-5 GRIDDED LIBRARY
82	X-6 SINGLE COPY PROBES
82	X-7 CONTIG ASSEMBLY
83	X-8 DATABASE ANALYSIS
83	X-9 X-GRAIL
83	X-10 GENOMIC CLONES AND LIBRARIES
83	X-11 EMBRYOS
84	X-12 cDNA CLONING, SEQUENCING AND DATABASE SEARCHES
84	X-13 WHOLE MOUNT IN SITU HYBRIDIZATION
84	X-14 RNA ISOLATION
84	X-15 NORTHERN ANALYSIS
84	X-16 RNASE PROTECTION ASSAY
85	X-17 RT-PCR DETECTION OF ALTERNATIVE HIRA mRNA
87	XI REFERENCES
113	XII ABBREVIATIONS - CONVENTIONS - DEFINITIONS - GLOSSARY - TERMS
117	XIII SUMMARY & SAMENVATTING
119	XIII-1 SUMMARY
120	XIII-2 SAMENVATTING
121	XIV CURRICULUM VITAE
125	XV LIST OF PUBLICATIONS
129	XVI DANKWOORD



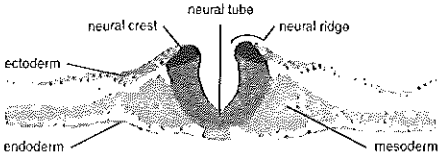
Introduction

I INTRODUCING THE INTRODUCTION

I 1 GENERAL

All vertebrate embryos start development as a fertilized egg that divides to form a multicellular clump of cells. Through directional cleavage and cell movement, from this forms an essentially three-layered structure of endoderm, mesoderm and ectoderm (the three germ layers) (**figure 1**). All adult tissues and organs form from these layers through processes that involve growth, migration, induction, differentiation and cell death. Ultimately, genes are at the basis of this development. A very complex sequence of genes being switched on, tuned and switched off at precisely defined points in the four-dimensional space-time continuum regulates the developmental process. Heissenberg, Turing (1953, 1990) and chaos theory suggest that there is an indeterminable stochastic aspect to it as well, but since that is very difficult to experimentally influence and study, for all practical purposes genes make up developmental biology. To paraphrase a well known saying: "Show me your genes and I will tell you what you are".

figure 1 Transverse section through a Hamburger and Hamilton stage 9 chicken embryo (posterior) showing the organization of the early embryo in three embryonic layers: endoderm, mesoderm and ectoderm. Also shows neurectoderm (neural tube and neural crest). Differential gray shading of various structures is artificial.



Pharyngeal arches (in literature and in this thesis also referred to as branchial arches) are, as the name implies, arches forming a "cage" around the pharynx. They come in left-right pairs starting around or below the lateral midline and meeting at the ventral midline of the embryonic head. Arches are visible as arches because of the external pharyngeal grooves or clefts. Internally, these grooves are matched by pharyngeal pouches (**figure 2**). This groove-pouch combination delineates the arches by forming constrictions in the mesoderm. The

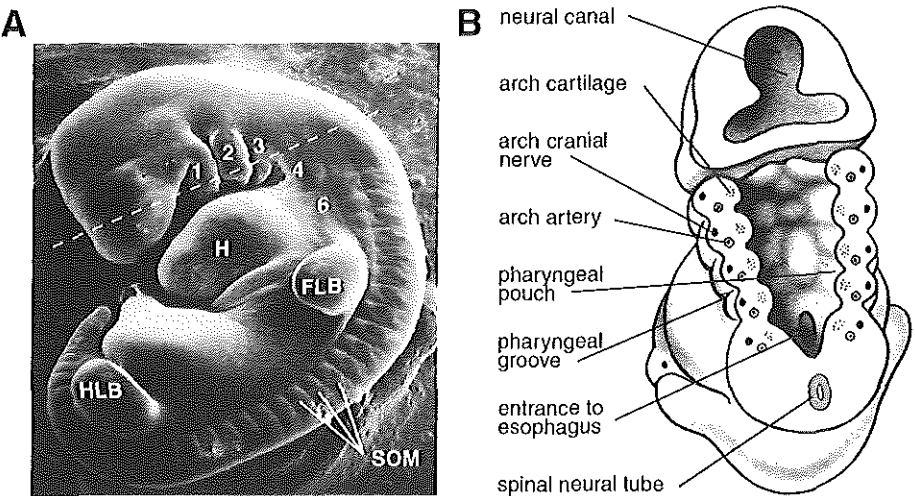


figure 2 A: A scanning electron microscope photo of an early fifth week human embryo highlighting the pharyngeal arches (numbered 1 through 4 and 6). Also clearly visible are the limb buds, heart and somites. Dashed line indicates the position of the cut resulting in the view depicted in B. B: Drawing showing the various features found in the pharyngeal region in an embryo dissected across the dashed line in A. Modified from Larsen (1993).

size of the arches diminishes from anterior to posterior; they are clearly visible for only a limited period because in higher vertebrates the second arch overgrows the posterior arches. For historical reasons the arches in vertebrates are numbered 1 through 4 and 6. Number five may be there as a separate entity very briefly, but that is a contentious issue. Branchial arches originate from the gill arches of fish. In fish gill slits separate the arches laterally because during embryonic development the grooves and pouches merge where in higher vertebrates the intervening tissue persists. Pharyngeal arches and pouches give rise to many components of the vertebrate neck and face and even some of the heart. Where there is such a broad spectrum of derivatives, there is the potential for a wide array of malformations. This is indeed the case with pharyngeal arches. Many of these anomalies are probably related to the cranial neural crest cells that populate the pharyngeal arches and contribute to the many arch derivatives. Neural crest constitutes a specialized subset of neurectodermal cells and is arguably one of the most interesting and important tissues in vertebrate development, and of vital importance to (amongst others) pharyngeal arch development. Congenital malformations can arise from external, i.e. epigenetic, disturbances, but very often there is a genetic defect at the basis of the syndrome. As genes are the foundation embryonic development is built on, study of the genetics of cranial neural crest related congenital malformations will

On HBA, chromosome 20 | | and C12C102. Ph.D. thesis by Larsen G. Writing describing efforts on the way about the molecular genetics of C12C102, a family of syndromes characterized by the maldevelopment or absence of (partially) cranial neural crest derived structures. See thesis, parasite.

provide us with answers to the question which genes are involved in neural crest and arch development. A link between (mal)development of the pharyngeal arches and a chromosomal defect is provided by a syndrome called CATCH22 or chromosome 22q11 Deletion Syndrome. CATCH22, featuring chromosome 22q11 deletions and malformation of arch derived structures, may prove to be the key to unlock the genetics of pharyngeal arch and cranial neural crest development.

I 2 SCOPE OF THIS THESIS

This thesis describes our efforts at analyzing the deleted 22q11 region and a candidate gene located therein. Firstly, I will show the genomic markers we mapped in the deletion. These markers enabled us to genetically analyze patients and assess the extent of their deletions, and provided a starting point for the construction of several cosmid contigs. I will further describe the subsequent detailed analysis of the genomic structure of the 22q11 region. By taking advantage of large scale sequencing efforts of the region I precisely mapped novel and established genes and determined their genomic structure. One of the genes in the region (more specifically its murine counterpart), *HIRA*, was analyzed in greater detail through cDNA analysis and expression studies in embryos and adults. Finally I will discuss the results in the context of the current state of knowledge.

II INTRODUCTION TO THE NEURAL CREST AND THE PHARYNGEAL ARCHES

II 1 GENERAL

Neural crest is a transient embryonic structure comprising of mobile cells that give rise to a wide variety of tissue types. The neural crest originates on the boundary between the neural plate and the epithelial ectoderm. Prospective neural crest is visible as neural ridges between the neurectoderm and the ectoderm (**figure 1**). Bringing neural plate in contact with ectoderm, *in vitro* or *in vivo*, mimics this process (Moury and Jacobson 1989, Dickinson *et al.* 1995, Selleck and Bronner-Fraser 1995). Crest precursors descend from both ectodermal and neurectodermal cells and progeny of induced neural crest cells can be found in neural crest, epidermis and neurectoderm (Bronner-Fraser 1988b, Bronner-Fraser 1995). Such experimentally induced neural crest yields actively migrating cells. As development progresses, the neural folds move closer to the dorsal midline, eventually fusing and enclosing the lumen of neural tube. Cells from the neural crest migrate away from their point of origin (the dorsal aspect of the neural tube) around the time of tube closure. In the timing of emergence relative to the time of tube closure the three classes of vertebrates where the neural crest has best been studied — *mammalia*, *aves* and *amphibia* — differ. In *amphibia* (*Axolotl*, *Xenopus*) the majority of crest cells leave the crest after tube closure. Only in the mesencephalon do they start migrating shortly before the neural folds fuse (Raven 1931). In mammalian (mouse) development cranial neural crest cells migrate before the tube has closed (Verwoerd and van Oostrom 1979, Nichols 1981, 1986). Avian (chicken, quail) cranial crest cells start migrating after tube closure (le Douarin 1983).

Evolutionary speaking neural crest cells are a relatively new cell type. Protochordates have an epidermal nerve plexus which could be the predecessor of both the neural crest and the epidermal



placodes, both of which are of ectodermal origin. Before there was a neural crest there was most likely a "pre-neural crest" in a common predecessor of the vertebrates and urochordates. The modern urochordate *Ciona intestinalis* expresses its homolog of a vertebrate neural crest marker in the dorsal neural tube but it does not have a neural crest (Corbo et al. 1997).

The evolutionary appearance of the neural crest closely correlates with the appearance of the vertebrate head. This head is thought to be a neomorph — an evolutionary new element — and not to be derived from existing head and trunk elements in invertebrates. Trunk and primitive head are segmented, so a new head derived from these structures would be expected to be segmented in a similar fashion. This is not the case: segmentation in the head, as I will discuss below, is organized in a different fashion. Somites are the main source of segmentation in the trunk because the neural tube is essentially unsegmented; any segmentation of neural tube derivatives (e.g. dorsal root ganglia) is imposed by the mesoderm. In contrast, the cranial neural tube is clearly segmented (initially in fore-, mid- and hindbrain (prosencephalon, mesencephalon, rhombencephalon)), whereas the morphological segmentation in the mesoderm is weak at best. Besides this difference in segmentation, an important difference between the new head and the trunk is the source of the bones, muscles and neurons. Trunk bones and muscles derive from the mesenchyme and trunk neurons from the neurectoderm. In the head region the neurectoderm, in the form of neural crest, contributes to all three of these structures. Almost all morphological differences between chordate and vertebrate heads are in structures that are derived from neural crest or dependent on it.

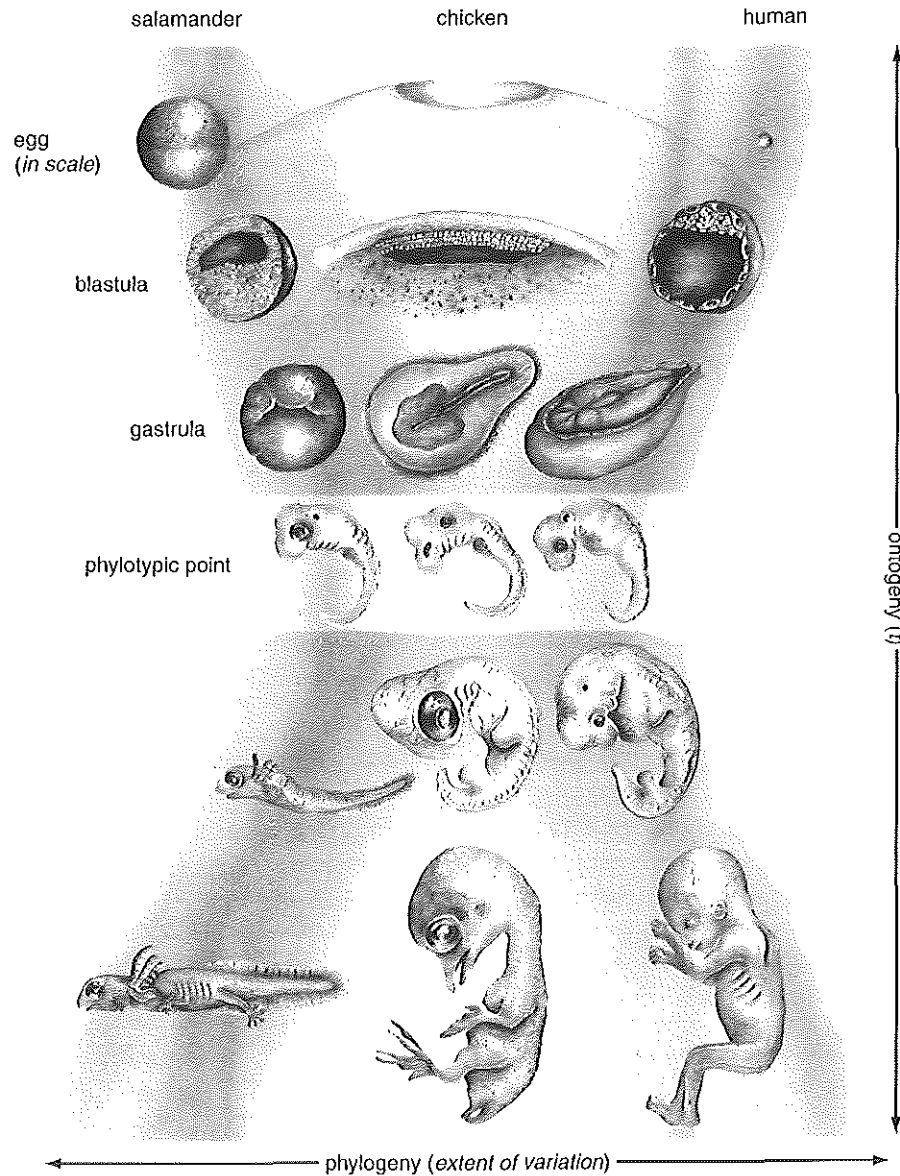
Given above mentioned segmentational and ontogenic novelties in the new head, it is reasonable to propose that the vertebrate head is a new feature and not a specialized trunk. Extra support comes from Kuratani, who describes the sharp migration boundaries between head and trunk neural crest, located at the level of the circumpharyngeal neural crest (Kuratani 1997).

Naturally, this new invention was brought about by a series of gradual changes. Gans and Northcutt propose the following

sequence of events (Gans and Northcutt 1983, Northcutt and Gans 1983, Gans 1988, Northcutt 1996). Early vertebrates were passive waterborne filter feeders with a distributed epidermal nervous system, diffusion gas exchange and a ciliated collagen-supported pharynx. This was followed by the development of a chorda and segmented muscle, allowing these protochordates to move in a more coordinated and active fashion. Increased movement would give selective advantage to a better gas exchange mechanism. The development of branchiomic muscles and flexible cartilaginous pharyngeal support (replacing the rigid collagen), combined with an improved circulation system (capillaries, muscled heart) would see to this. This could be considered the start of the development of the new head. Next it is suggested that these protovertebrates developed predation as a means of obtaining food. This would involve the development of paired external mechano-chemical sense organs and a centralized nervous system (brain) to cope with the complex coordinated processing of sensory input and muscle actuation. The agnatha developed patches of dermal dentine as electroreceptors (used for predation and avoiding predation). Later these became surrounded and supported by dermal bone. Then jaws developed, facilitating the capture and ingestion of larger prey. A larger body size, concomitant with the advantage of capturing even larger prey, necessitated the development of a support system (vertebrae): the vertebrate was born.

All vertebrates, from fish to frogs to fowl, are built on the same basic plan where the pharyngeal arches figure prominently. It has been suggested that the aquatic ancestry of all vertebrates is still visible in the vertebrate embryo. Haeckel drew a famous set of pictures of embryos of various vertebrate species to show their high degree of similarity. At some stage during development vertebrate embryos look roughly the same, whereas they started as differently shaped zygotes and eventually turn into very different animals. This is referred to as a phylogenetic egg-timer where there is a broad range of start and end points, but a very narrow pass-through at the phylotypic stage (figure 3). Recent research has shown that the Haeckel drawings are highly idealized (Pennisi 1997, Richardson 1997), but the fact remains that for their hugely different final phenotypes (compare frog, eagle and ray), there is a window during development where their embryos are remarkably similar, suggesting there is some type of constraint on the permissible variation. Due to heterochrony and allometric growth, there is quite a variation in the timing of the development of organs and other features. For example, the formation of teeth, the descendants of the early electroreceptors, involves an elaborate cascade of bidirectional induction processes between neural crest and ectoderm. Such a cascade lends itself exquisitely to variation with respect to timing and spatial arrangement. This is thought to be the driving force behind the emergence of variation in neural crest derived structures between vertebrate species (Hall 1990). Assuming the lamprey is the most primitive vertebrate today, one sees that the timing of these interactions is shifted in mammals and birds: mammals shifted to a later time point, birds to an earlier. With respect to the

affected structure. It also deals with the molecular genetics of other regional malformations associated with chromosomal rearrangements, multiple genes or their genes, and the delineation of single gene anomalies and congenital gene syndromes. Naturally the clinical aspects of the CATCH22 syndrome.



with the various neural crest cells with somata, cranial innervation and motor efference, are discussed. The neural crest is an embryonic structure, some call it the fourth germ layer, generated when surface ectoderm and mesoderm meet. Neural crest cells are highly mobile and will sort

cephalic neural crest as a source of variation, Irmak and Ozcan (1997) make the interesting observation that environmental adaptations of humans seem to be limited to neural crest derived structures (pigmentation being the most obvious one).

11 2 THE RHOMBENCEPHALON IS SEGMENTED INTO ODD AND EVEN NUMBERED RHOMBOMERES WITH SEPARATE IDENTITIES

Above I alluded to the initial segmentation of the cranial neural tube into the three segments fore-, mid- and hindbrain. Later in development these segments are divided into subsegments; rhombencephalon subsegments are called rhombomeres, which are visible bulges dividing the hindbrain in eight domains (figure 4). Rhombomeres play an important role in defining subsets of neural crest cells. Migrating crest cells follow distinct pathways, depending on the level and timing of emergence on the antero-posterior axis. The neural crest cell's ultimate fate is intimately linked with the migratory pathway the cell has chosen (see later). The two major pathways are the dorsolateral route (underneath the epidermis) and the ventral route (close to the neural tube). In the head the emergence of distinct streams of crest cells on the rostro-caudal axis is partly related to the apparent lack of neural crest cells originating from the third and fifth rhombomere. Crest cells migrating from rhombomeres to pharyngeal arches form distinct streams establishing a relationship between the rhombomeres and the pharyngeal arches (figure 5). Roughly, the crest from rhombomeres 1 and 2 migrates to arch 1 and gives rise to derivatives associated with this arch; rhombomeres 3 and 4 do the same for arch 2, while rhombomeres 5 and 6 do so for arch 3. Originally it was thought that rhombomeres 3 and 5 did not produce neural crest cells, but labelling experiments show that they actually do, be it not as many as

figure 3 Phylotypic egg-timer. While both the start and end points of vertebrate embryonic development are widely variable (top and bottom), all embryos pass through a phylotypic point where the embryos are remarkably similar (narrow center), possibly indicating a basic molecular-genetic mechanism shared among all vertebrates at this stage.
 After Larsen (1993) and Duboule (1994).

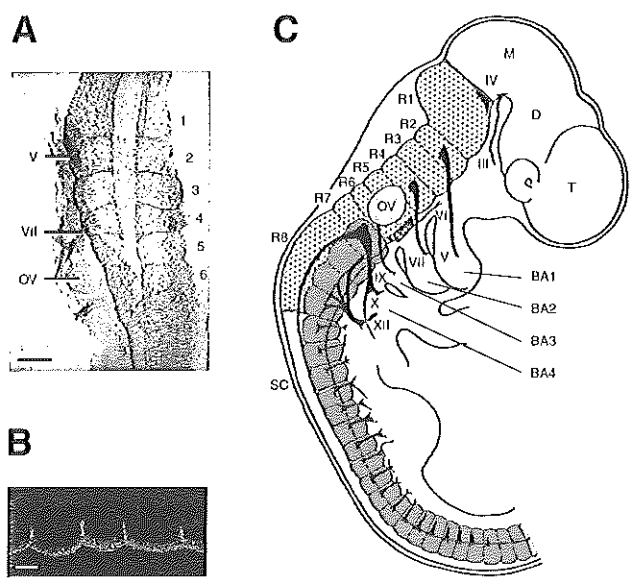


figure 4 **A:** Photomicrograph of a wholemount 2 day old chick embryo hindbrain stained with anti-68 kD neurofilament antibody. Rhombomeres are labeled 1 through 6. OV: Otic Vesicle; V: fifth (trigeminal) cranial nerve root; VII: seventh (facial) cranial nerve root. Scale bar equals 300 μm . **B:** Parasagittal section of a 3 day old chick embryo hindbrain stained by immunofluorescence with anti-68 kD neurofilament antibody clearly showing rhombomere boundaries. Scale bar equals 100 μm . **C:** Schematic drawing of a Hamburger and Hamilton stage 18 chicken embryo showing the hindbrain and its rhombomeres (medium gray, labelled R1 through R8), the cranial motor nerves (III-XII), and branchial arches (BA1-BA4). The mesodermal somites (dark gray) and spinal motor nerves are also shown; somites alongside R7 and R8 are dispersed (broken lines). OV: Otic Vesicle; M: Metencephalon; D: Diencephalon; T: Telencephalon; SC: Spinal Cord. Modified from Lumsden and Keynes (1989) (A, B) and Keynes and Lumsden (1990) (C).

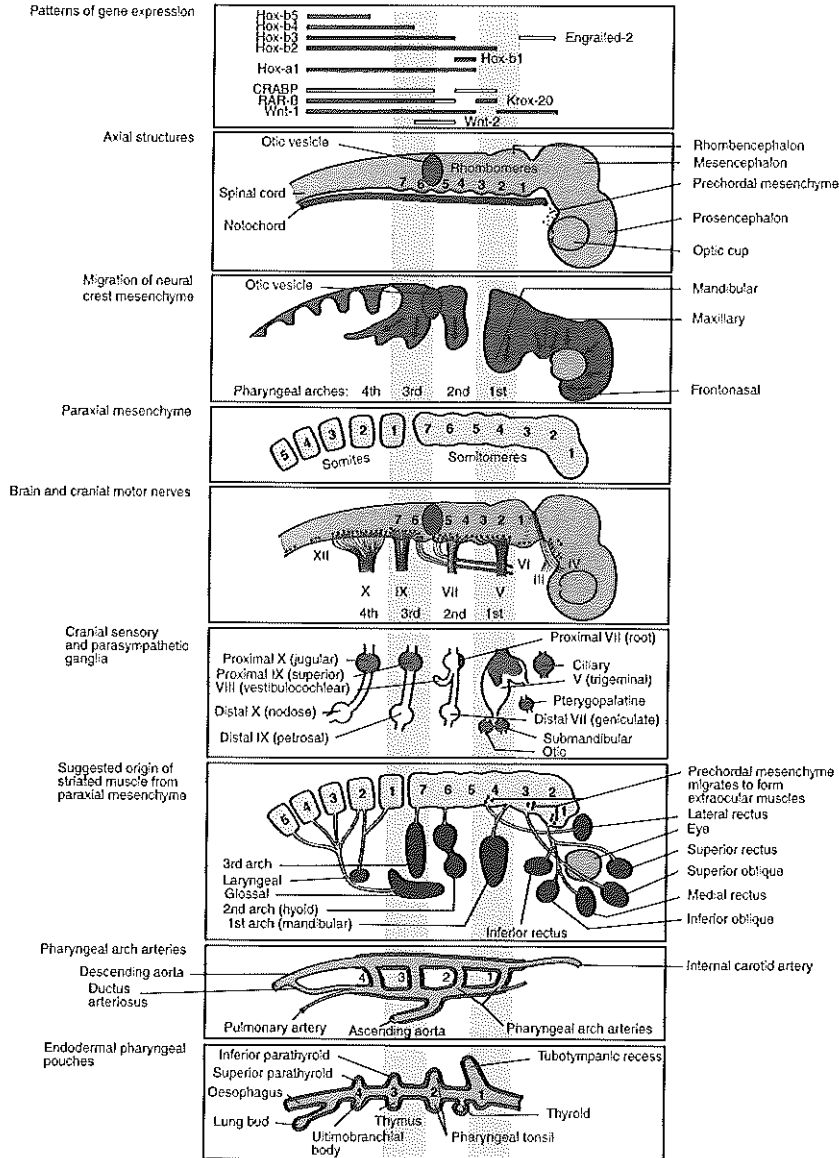
the other rhombomeres. Many of the rhombomere 3 and 5 crest cells die by apoptosis (Jeffs and Osmond 1992, Jeffs *et al.* 1992), induced by their neighbors (Graham *et al.* 1993). Isolated third and fifth rhombomeres give rise to many migrating crest cells *in vitro*. But when contact with their even-numbered neighbors is re-established this stops (Graham *et al.* 1994, 1996). In stead of adding neighbor rhombomeres one can also add proteins normally produced in these rhombomeres. This induces a sequence of genetic events that restores the cell death program (Graham *et al.* 1996, Marazzi *et al.* 1997).

Migrating from their points of origin to their final destination around the inner of neural tube. They mostly travel either just under the ventricles, or closer to the neural tube through the astrocyte spaces. [CH] or astrocytic many seen an important role in the regulation of neural crest cell migration.

Those crest cells in the third and fifth rhombomere that do not succumb to apoptosis migrate rostrally and caudally to exit the neural tube at the neighboring rhombomeres (2 and 4, and 4 and 6 respectively), joining the streams of cells emanating from these. Rhombomere 5 yields a segregated population of crest cells separated on whether or not they express molecular markers like *Krox20*. *Krox20* positive cells migrate caudally to join the rhombomere 6 cells populating branchial arch 3; negative cells move anteriorly and join the rhombomere 4 cells to populate arch 2 (Nieto *et al.* 1995). Receptor Kinase family proteins have been implicated in maintaining the different identities of odd and even rhombomeres and keeping the streams of migrating crest cells separate. (Becker *et al.* 1994, Xu *et al.* 1995). The key to the odd-even identities may lie in the extracellular matrix. Cells from odd numbered rhombomeres mix more easily with each other than with cells from even-numbered rhombomeres, and *vice versa* (Guthrie *et al.* 1993). Transplantations and rotations in chick embryos resulting in abutting even-even or odd-odd rhombomeres (minus their boundaries) often cause fusion of rhombomeres with equal parity; a new boundary is generated where odd-even rhombomeres come in contact (Guthrie and Lumsden 1991, Lumsden and Guthrie 1991). Under normal circumstances cells crossing over from odd to even rhombomeres will take the identity of the new rhombomere. In the even numbered rhombomeres a signalling pathway could exist that induces a cascade of actions leading to the identity switch. At boundaries special conditions would have to apply: maybe a cell needs a certain above-threshold number of neighbor cells interacting with it before a switch takes place.

In chicken embryos rhombomere boundaries are morphologically distinct. Cell density is lower and cells are distributed towards the ventricular (luminal) side, leaving space basally for an extra cellular matrix different from that in the rhombomere bodies (Lumsden and Keynes 1989, Layer and Alber 1990, Guthrie *et al.* 1991, Heyman *et al.* 1993). These inter-rhombomeric borders may act as barriers for planar signals (Martinez *et al.* 1995), just as they function as virtually impenetrable clonal segment borders. Descendants of rhombomeric cells labeled before border formation can be found across borders, but the majority of cells labeled after border formation only yield progeny within the rhombomere (Fraser *et al.* 1990). The barrier is not perfect though; some (but very few) clones do cross boundaries (Birgbauer and Fraser 1994).

When odd and even numbered rhombomeres are discussed, in general only rhombomeres 2 through 6 are considered. Rhombomeres 1 and 7 and 8 are considered special cases: the boundary between rhombomere 7 and 8 is not very clear, and rhombomere 1 and 7/8 form the boundaries with neighboring neural tube segments, which in several cases causes them to behave differently from the other five rhombomeres. Moreover, r7 and r8 overlap with the start of clear mesodermal segmentation (somites) (figure 4C), so in that region we find both neuroectodermal and mesodermal segmentation with all the interactions (conflicts?) between segmentational signals that



Many (CN) or related components such as trigeminal, ciliary and submandibular have a profound influence on the migration of neural crest cells. There are also many growth factors that regulate where and when crest cells migrate. In a very brief manner, neural crest cells within the neuronal

brings with it. I only highlighted a few examples of difference between odd and even rhombomeres; more examples are discussed in Lumsden and Keynes (1989).

II 3 FATES, DESTINATIONS AND CONTRIBUTIONS OF NEURAL CREST

II 3 1 NEURAL CREST CELLS ARE MULTIPOTENT: SUMMARY OF IN VITRO RESULTS

Neural crest cells give rise to a large number of derivatives, such as neural (cranial sensory neurons, enteric neurons, sympathetic and parasympathetic neurons), neuroendocrine (adrenal medulla), neural support (glia cells), mesodermal (craniofacial bone, cartilage, skeletal muscle connective and smooth muscle tissue, thymic, thyroid, parathyroid and cardiac tissue) and ectodermal (melanocytes) cell types.

Generally it is assumed that a given neural crest population consists of a mix of predetermined and pluripotent cells. See **figure 6A** for the results of *in vitro* clonal analysis of quail cephalic neural crest by **Barroffio et al. (1988)**. Studying the clonal progeny of cultured cells shows that in early neural crest there is a neuronal-fate restricted population and that glial and melanocyte restricted populations arise later. The initial population contains both pluripotent precursors and precursors that give rise to single phenotype colonies (**Henion and Weston 1997**). Crest cells that have already migrated into the branchial arches still consist of a heterogeneous population. Mesenchymal cells from posterior branchial arches grown *in vitro* for clonal analysis (using cell type specific immunohistochemical and morphological markers) can be divided in four different groups. Each group yields different combinations of cell types (ectomesenchymal, neuronal, muscle, connective tissue), the relative proportions changing with progression of time (**Itoh and Sieber-Blum 1993**).

figure 5 Schema illustrating the organization of the head and pharynx in an embryo at about stage 14. The individual tissues have been separated but are aligned in register through the shaded zones (representing the pharyngeal arches). from Gray (1993).

Notwithstanding the pluripotent nature of neural crest cells, there is regional variation in potentiality. Between different cranial neural crest populations of the mouse there is a difference in the potentiality to develop certain fates (Chareonvit *et al.* 1997) (figure 6B). Neural plate explants from different anterior-posterior locations (fore- or midbrain), give rise to different spectra of neural crest cell derivatives; where melanocytes are formed from forebrain explants, midbrain explants do not give rise to melanocytes. Conversely, collagen type II expressing mesenchymal type cells arise from midbrain cultures, but not from forebrain. Between trunk and cephalic neural crest there is a difference in fate restrictions as well. And there are indications that even within the trunk there is a difference in the distribution of fates, both in space (Asamoto 1995) and time (Reedy *et al.* 1998).

11 2 FATE MAPPING SHOWS THE MYRIAD CONTRIBUTIONS OF THE NEURAL CREST

In vivo the temporo-spatial coordinate of a given neural crest cell population determines the spectrum of cell types in its progeny. The actual number of different derivatives the progeny gives rise to is generally smaller than the potential number. The potentiality of migratory and pre-migratory neural crest cells has been subject to many experiments showing the multipotency and heterogeneity of crest populations (Bronner-Fraser *et al.* 1980, Sieber-Blum and Cohen 1980, Ciment and Weston 1985, Girdlestone and Weston 1985, Barbu *et al.* 1986, Barald 1988, Barroffio *et al.* 1988, 1991, Ito and Sieber-Blum 1991; see also previous chapter). Heterochronic and isochronic transplantations of quail mesencephalic neural crest to chicken show that early and late migrating cells have equal developmental potential (Baker *et al.* 1997). An early migrating population normally migrates both dorsally and ventrally, while late migrating cells are confined to the dorsal pathway. Heterochronic transplantations (substituting late neural crest for early or vice versa) results in the donor tissue behaving according to the timing in the host with regard to pathway choice: late donor cells now migrate ventrally (which they normally would not) to contribute to jawbones and cartilage when taking the place of

each, they have not shed their neural crest background completely. The epithelial and epithelial-mesenchymal transition genes are regulated by their molecular, allowing the formation of avian and mammalian, respectively. Not surprisingly, laminar protein, which

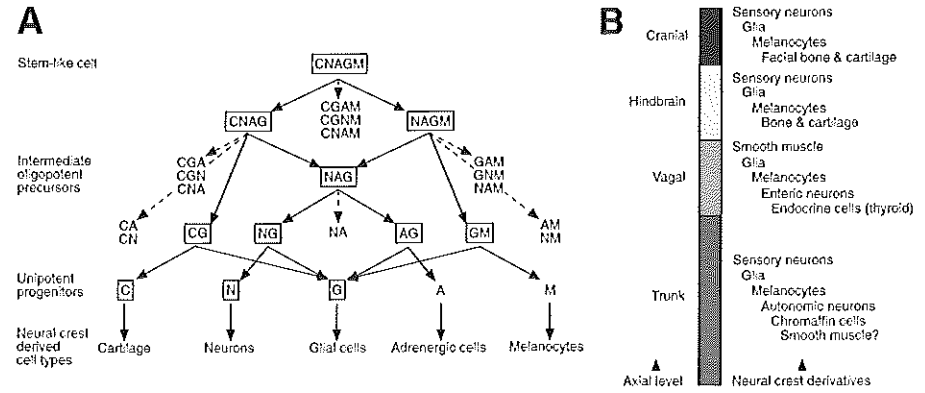


figure 6 Neural crest derivatives. **A:** Diagram illustrating all types of progenitors from quail cephalic neural crest cells as revealed by phenotypic analysis. Non-framed progeny is putative, framed is actually observed by Barroffio *et al.* (see references in le Douarin *et al.* (1994)). These results suggest an increased specialization of crest cells through time, from a totipotent neural crest stem-like cell to unipotent progenitors via intermediate oligopotent precursors. C: Cartilage; N: Neurons; G: Glial cells; A: Adrenergic cells; M: Melanocytes. **B:** Variations in avian neural crest derivatives produced at different rostrocaudal levels of the neuraxis. After Le Douarin *et al.* (1994) (A) and Anderson (1997) (B).

early cells. Early donor cells restrict to the dorsal path when taking the place of late host cells. When late cells are transplanted to an ablated late host (where the neural crest was removed to prevent crest cell migration), donor cells migrate ventrally. This shows that late emerging neural crest cells mainly take the dorsal path because early migrating cells have already occupied the ventral space. It seems that, in the cranial region at least, opportunism rather than predetermination is what sets neural crest cells off on their different pathways.

In the trunk there is little regionalization along the anterior-posterior axis: neural crest cells contribute neurons and support cells to ganglia and give rise to melanocytes. Contribution to the adrenal medulla, however, is level dependent in the chick (le Douarin and Teillet 1973), and some specialized structures like fins receive contributions from localized trunk crest (Collazo *et al.* 1993). In contrast, in the head region there is distinct fate restriction. As discussed above, crest cell populations are generally multipotent (with some fate restriction) when given free reign, but *in vivo* the spectrum of derivatives is more restricted. Also, the morphological appearance of similar derivatives varies with the anterior-posterior axis. Crest cells migrating to the anterior branchial arches may all form bones, cartilage and ganglia, but the morphology of these derivatives is characteristic for each arch (compare hyoid with lower jaw) (figure 7). Different regions in the hindbrain have been given specific names to indicate the fate of the corresponding neural crest. Hence we have the vagal crest (named after the nervus vagus, number X), divided in cardiac crest (rhombomeres 6, 7 and anterior 8, up to and including somite 3 level) and enteric crest (posterior rhombomere 8, next to somites

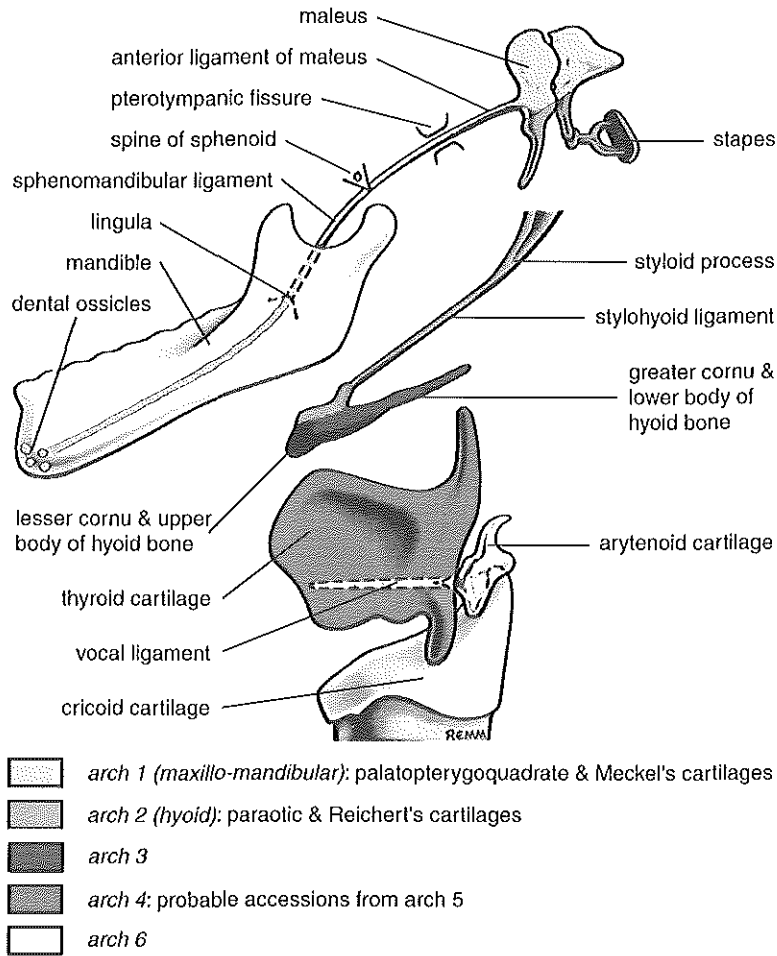


figure 7 Schema illustrating the skeletal derivatives (osseous and cartilaginous) of the pharyngeal arches (viscerocranium). Modified from Gray (1995).

4 and 5), as subdivisions of the hindbrain crest. See **figure 4C** for orientation.

Through various techniques researchers have studied which tissues and structures are derived from the neural crest. Craniofacial bone and cartilage formation have been studied extensively, mainly in chicken and quail. One of the techniques used for this is cell labelling. Individual cells or groups of cells can be labeled with fluorescent dyes like Dil and LRD (Lysinated Rhodamine Dextran) *in vivo*. Cells can be injected individually or the dye can be injected in for example the lumen of the neural tube, after which cells take it up. Either technique allows the tracing of the lineage of the labeled cells. Another often used technique is transplantation of fluorescent (or, in the past, radioactively) labeled chicken cells or labeled or unlabeled quail cells into chicken hosts. Quail cells can be histologically distinguished from chicken cells thanks to their unique nuclear morphology. A third technique is immunohistochemistry or *in situ* histochemistry with neural crest specific markers like the HNK-1 antibody (Vincent *et al.* 1983) or *slug* ribonucleotide probes. HNK-1 recognizes a carbohydrate epitope originally described for human natural killer cells (hence the name) (Abo and Balch 1981). It is found on many (not all) migrating neural crest cells, but also in the perinotochord and in other locations (Kruse *et al.* 1984).

3 CRANIAL NEURAL CREST AND THE FIRST TWO BRANCHIAL ARCHES PRODUCE MOST OF THE FACE

Derivatives of cranial neural crest cells form most of the facial bones and muscle connective tissue and ganglia in chicken and mouse (le Lièvre and le Douarin 1975, Noden 1983a, b, 1984, 1988, Couly *et al.* 1993). Avian cranial sensory ganglia derive from both the cranial neural crest and neural placodes (Noden 1988) (neural placodes are patches of the neurectoderm that "stayed behind" in the ectoderm and only migrate inwards on interaction with neural crest cells). In chicken a very limited contribution of neural crest cells to some facial muscles has been shown by le Lièvre and le Douarin, and the connective tissue for some of the avian facial muscles is neural crest derived (le Lièvre and le Douarin 1975). Chicken

chondroid (cartilage) tissue is of neural crest origin, while secondary cartilage is of crest or cephalic mesoderm origin, depending on the location (Lengele et al. 1996).

Each branchial arch gives rise to a unique morphology of skeletal, connective and muscle tissues in higher vertebrates. Neural crest cells migrate into the branchial arches and the frontonasal area as ectomesenchyme. The first arch is a C shaped arch arching the stomodeum (prospective mouth). Maxillae (upper jaw) will form from the dorsal or rostral end (which is close to the frontonasal prominence), while the ventral part of the first arch (abutting the second arch) is responsible for the formation of the mandibles (lower jaw). Meckel's cartilage is the primordium of the lower jaw and extends up to the otic capsule. (figure 7) The dorsal end close to the otic capsule becomes separated to form the malleus of the inner ear. Possibly it also forms the inner ear incus, but this may also be derived from the maxillary part of the first arch cartilage (the palatopterygoquadrate cartilage). During further development of the embryo, the initial cartilages of the first arch become encased in bone or are absorbed or undergo secondary ossification. Secondary ossification is a process where the initial bone is laid out as cartilage that later undergoes ossification, as opposed to bones that form from direct ossification of mesenchyme. Ossification, and possibly also chondrogenesis, in the head requires mesenchymal-ectodermal interactions (Noden 1978, Hall 1978a, b, 1980, Bee and Thorogood 1980).

Reichert's cartilage is formed by the second or hyoid arch, which, as the name suggests, is responsible for the formation of the hyoid (although the third arch also contributes) (figure 7). Part of the inner ear; the stapes, originates from Reichert's cartilage as well (figure 7). The

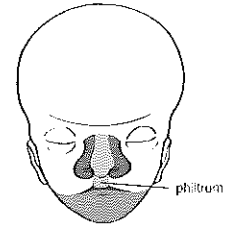


figure 8 Drawing showing the parts of the face that share common embryonic origins. Philtrum and medial part of the nose derive from the two medial nasal processes (which later fuse to form the intermaxillary process) (medium-light gray). Cheeks and upper lips and jaw derive from the maxillary swellings (light gray). The lateral nasal processes give rise to the lateral parts of the nose (dark gray) and the mandibular swelling to the chin and lower jaw (medium-dark gray). From Larsen (1993).

transplanted tissue and host tissue. Contribution of the neural crest varies depending on whether the tissue was generated and when. Cranial neural crest differentiates basal from trunk neural crest which form muscle, endoderm and mesoderm support cells by forming or contributing to tissue.

second and third arches contribute to various other parts of the inner ear. Once again we see neural crest contributions to the inner ear. The outer ear is formed by the dorsal ectoderm of the first and second arches, i.e. the area surrounding the dorsal part of the hyomandibular (first pharyngeal) groove, and by mesenchyme from the same two branchial arches. Neural crest may contribute directly and indirectly (through mesenchymal-ectodermal interactions) to the formation of the outer ear.

Palate development consists of a the formation of a primary palate and a secondary palate. The primary palate forms from mesencephalic neural crest of the frontonasal mass and forms the palate in the anterior part of the mouth cavity. Essentially the same neural crest population gives rise to the frontonasal prominence that interacts with the maxillary prominences to form the philtrum and nose (figure 8). Posterior of the primary palate the palate forms from two halves constituting the secondary palate. Two palatal shelves arise from the bilateral maxillary processes and grow vertically down the sides of the tongue. At a precise developmental time point these shelves rise to a horizontal position above the tongue and continue growing until their medial edges come in contact (figure 9). In mammalia the midline epithelial cells then either die through apoptosis or migrate to the epithelium of the oral and nasal cavities so that the mesenchyme becomes continuous and the palate closed. In birds and some reptiles the situation is different in that the palatal shelves grow horizontally *ab initio*, that the shelves make contact but never fuse and that the medial epithelium keratinizes in stead of dies. This leaves these animals with a natural cleft palate (in some reptiles the epithelia do not even make contact, resulting in a large choanal groove). Other reptiles and also amphibians have mainly a primary palate. Contact between the shelves is not necessary for the changes in the midline epithelium to occur: isolated halves in culture still display the epithelial differentiation (apoptosis, migration or keratinization) characteristic for the respective species *in vivo* (Ferguson et al. 1984). In

general the specification of the particular nature of the midline epithelium originates from the mesenchyme (which is neural crest derived), as was shown by heterotopic, heterochronic and hetero-specific epithelial or mesenchymal transplantations (Ferguson and Honig 1984). Secondary palate development

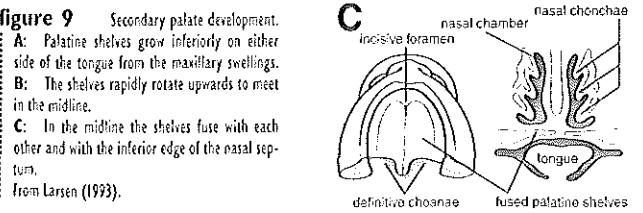
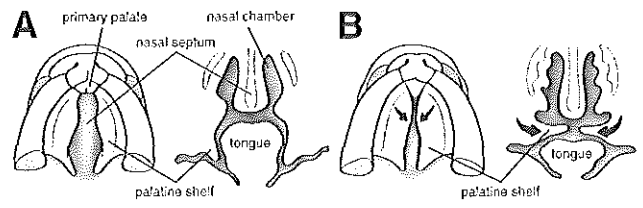


figure 9 Secondary palate development. A: Palatine shelves grow inferiorly on either side of the tongue from the maxillary swellings. B: The shelves rapidly rotate upwards to meet in the midline. C: In the midline the shelves fuse with each other and with the inferior edge of the nasal septum. From Larsen (1993).

involves a host of molecules regulating epithelial-mesenchymal interactions and structural-mechanical behavior. An important role is played by Extra Cellular Matrix (ECM) molecules such as collagens, proteoglycans and hyaluronic acid. See a later chapter for a general discussion on ECM and neural crest. A secondary palate appears to be a feature of higher vertebrates, as further down the evolutionary ladder (fish) the secondary palate is absent (Ferguson 1988).

It has recently been shown that there is a one-to-one mapping of the axial level of origin of facial skeleton and facial musculature. Rhombomere 1 and 2, rhombomere 4 and rhombomere 6 and 7 crest form bones with sharp boundaries corresponding to the rhombomere(s) of origin, but these boundaries do not coincide with anatomical landmarks. Connective tissue of skeletal muscle is derived from neural crest cells of the same origin as the bone it is attached to. This is true for branchial and hypoglossal (tongue) muscles as well as for the other craniofacial muscles (Kontges and Lumsden 1996). The muscles themselves originate from mesoderm of the same axial level as the crest cells that formed the bones they attach to. The neural crest and paraxial mesoderm cells migrating to their final destinations to generate the muscles and bones mix quite extensively in most of the head region. Only in the branchial arch region there is a segregation of neural crest and mesodermal populations (Trainor and Tam 1995).

The observed distinct functional segmentation of the preotic head mesoderm — paralleling neural segmentation — does not correspond to a morphological segmentation like in the trunk, where the mesoderm is segmented in the form of somites. Claims of the existence of so called somitomers in the head have been made, based on shallow transverse furrows dividing the mesoderm (Meier 1981, Meier and Tam 1982, Gilland and Baker 1993), but the exact significance is unclear. Regionalization of the mesoderm in the head is shown by Noden (1986) to be a neural crest affair: trunk somites replacing somitomers form regionally correct muscles, indicating that the crest derived cranial mesenchyme communicates positional cues to presumptive myoblasts. Regionalization paralleling the neurectoderm in the hindbrain is also seen in the ectoderm by heterospecific isotopic transplantations. So-called ectomers with a one-to-one relationship to specific hindbrain regions were observed by Couly and le Douarin (1990).

Noden (1983a) performed some classic transplantation experiments of presumptive 1st branchial arch neural crest cells to the position of presumptive 2nd or 3rd branchial arch level in ablated embryos. The cells migrated properly for their position — i.e. to 2nd or 3rd branchial arch — but developed more or less like they would originally have, i.e. like 1st branchial arch. At the stage the transplantations were performed the regional morphogenetic specification of the hindbrain neural crest thus seems to be determined and not influenced by the pharynx, pharyngeal pouches or ectoderm. Identity is also preserved when craniofacial neural crest is transplanted between species: donor cells form donor species specific structures in their host. (Andres 1946, 1949, Wagner, 1949, 1959). Clearly, the choice of destination and final (morphological) differentiation are two

independent choices. The general tendency is that ectopic crest cells migrate in a host specific manner, as shown by the rhombomere 1 transplantations and by rotating rhombomere pairs anterior-posteriorly (Sechrist *et al.* 1994). Early determination of morphogenetic fate does not seem to occur in the case of derivatives such as enteric ganglia and melanocytes, because heterotopic transplantation results in donor cells contributing to these structures in a host specific fashion (Noden 1983a).

4 THYMUS, THYROID AND PARATHYROIDS DEPEND ON NEURAL CREST FOR PROPER DEVELOPMENT

Thymus and parathyroids are neural crest cell induced third and fourth pharyngeal pouch derivatives. Later in development they move to more posterior positions to find their final position around the trachea and esophagus (figure 10). Two thymus progenitors arise from the ventral aspect of the third pharyngeal pouch. Proper thymus development involves a precisely timed sequence of migration and interaction of cells from different sources. Mesenchyme and endoderm are not the only tissues involved: blood-borne lymphoid stem cells populate the thymus where they proliferate and differentiate (Auerbach 1960, Moore and Owen 1967, le Douarin and Jotereau, 1975). Not only is neural crest derived mesenchyme necessary for induction, it also contributes structurally to the thymus delivering its mesenchymal components, i.e. the septa between the lobules and the sheaths around the blood vessels (le Lièvre and le Douarin 1975). Neural crest cells also line and penetrate the endodermal tubes that form the thymic primordia. There they control the proliferation and differentiation of the thymocytes. A bilobate thymus gland is formed when the two primordia fuse after migrating down to a location ventrocaudal of the thyroid gland. Direct involvement of the neural crest has been proved by the observed thymus hypoplasia or aplasia in chicken embryos with partially ablated cranial neural crest (Bockman and Kirby 1984, 1985, Kuratani and Bockman 1990).

Parathyroids produce a hormone (Parathyroid Hormone PTH) that regulates systemic calcium levels. The four parathyroids, two inferior and two superior, arise from the dorsal portion of the third pharyngeal pouch and from the fourth pouch, respectively (figure 10). Similar to the thymus the mesenchymal components of the parathyroids originate from the ectomesenchyme, and like the thymus, endodermal rudiments detach from the pharynx wall to migrate caudally. The thyroid gland to which the parathyroid glands attach themselves, also receives neural crest contributions: its C cells, which produce PTH-counteracting Calcitonin, are neural crest derivatives. Thyroid development starts as an invagination of the pharyngeal endoderm behind the prospective tongue, around the level of the first pharyngeal pouch (figure 10). This organ too migrates caudally to a position around the trachea. As is the case with the thymus, deletion of cranial neural crest results in absence or improper development of thyroid and parathyroids (Bockman and Kirby 1984).

5 REMODELING OF THE PHARYNGEAL ARCH ARTERY SYSTEM IS CRUCIAL FOR ANTERIOR CARDIOVASCULAR DEVELOPMENT

As complex as the heart is in the developed vertebrate, development starts out with two simple tubes, initially far apart on the very lateral edges of the embryo proper. When the presumptive gut closes to form the foregut, the two endocardial primordia come in close contact and eventually fuse to form the endocardial tube. Through a complex sequence of events, the heart tube develops local swellings and folds several times to form the compact multi-chambered heart we know (figure 11) (Fishman and Chien 1997). While structurally the bulk of the heart originates from the cardiac endo- and mesoderm, pharyngeal arches 3, 4 and 6 are of vital importance for proper development. At a certain stage of heart development, the aorta and pulmonary artery have a common trunk (the conotruncus) attached to the presumptive right ventricle. Cells from the so-called cardiac neural crest migrate through the three

Note as the heart develops, the endocardial tube is remodeled into a complex network of vessels. The endocardial tube is remodeled into a complex network of vessels. The endocardial tube is remodeled into a complex network of vessels. The endocardial tube is remodeled into a complex network of vessels.

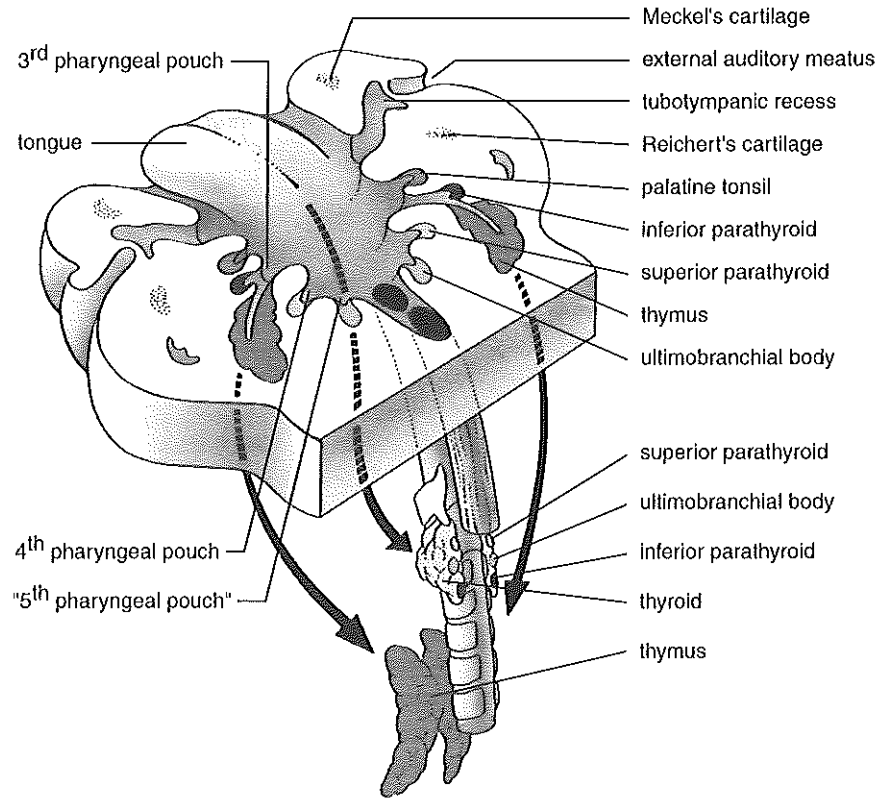


figure 10 Development of the pharyngeal pouch derivatives. All of the pharyngeal pouches give rise to adult structures. Pouch 1: tubotympanic recess; pouch 2: palatine tonsils; pouch 3: inferior parathyroids and thymus; pouch 4: superior parathyroids and (though possibly from a hypothetical pouch 5) ultimobranchial (telopharyngeal) bodies. The thyroid forms from an unpaired primordium behind the tongue. Thyroid, parathyroids, thymus and ultimobranchial bodies separate from the lining of the pharynx and migrate to their definitive locations in the neck and thorax. Modified from Larsen (1993).

posterior branchial arches and on towards the heart to separate the two great vessels by forming the outflow septum (aorticopulmonary septum) (Rychter 1978, Thompson and Fitzharris 1979, Kirby *et al.* 1983). Aorticopulmonary septation is intimately linked to ventricular septation. The ventricular septum starts caudally and grows rostrally to meet the aorticopulmonary septum in the conal region (the conus is the proximal part of the conotruncus). This process separates the left ventricle from the right ventricle and connects the pulmonary trunk to the right ventricle and the

aorta to the left, each with their own valve. The ventricular septum has recently been confirmed to be of neural crest origin (Waldo *et al.* 1998). Neural crest contributions are also found during the ongoing remodeling processes that take place after septation. The semilunar valves in the aorta and pulmonary are of neural crest origin, as is the tunica media of the proximal walls of the great vessels. Later in development the neural crest cells in the outflow tract are eliminated through apoptosis and replaced with cardiomyocytes (Poelman *et al.* 1998). Cardiac Ganglia are neural crest derived (Kirby and Stewart 1983) and cardiac neural crest cells are found in various other minor areas (Waldo *et al.* 1998). In these locations the crest cells possibly contribute in a way other than structurally. For example, some of the coronary arteries are dependent on neural crest cells, although crest cells do not contribute to the arteries directly (Hood and Rosenquist 1992, Waldo *et al.* 1994) and abnormal differentiation and function of the myocardium during early heart development following cardiac neural crest ablation suggest some novel regulatory role for neural crest cells (Waldo *et al.* 1999).

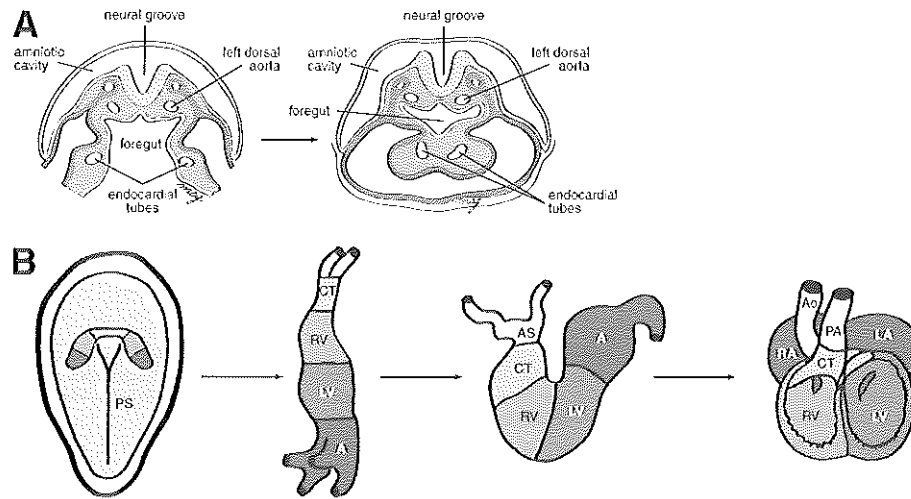


figure 11 Cardiac Morphogenesis. **A:** In the earliest stages of heart development the two endocardial tubes are located in the ventral mesoderm at the lateral edges of the embryo. Cephalocaudal and lateral folding brings the two tubes into the ventral midline, where the tubes will fuse over most of their length to form the primary heart tube. **B:** The region of the embryo that will eventually form the heart, the precardiogenic mesoderm, is specified at the primitive streak stage. A fused and straight primary heart tube is specified to form various regions and undergoes extensive looping and modifications to form to multichambered mature heart. PS: Primitive Streak; CT: Conotruncus; RV: Right Ventricle; LV: Left Ventricle; AS: Aortic Sac; A: Atria; RA: Right Atrium; LA: Left Atrium; Ao: Aorta; PA: Pulmonary Artery.

Modified from Larsen (1993) (A) and Olson and Srivastava (1996) (B).

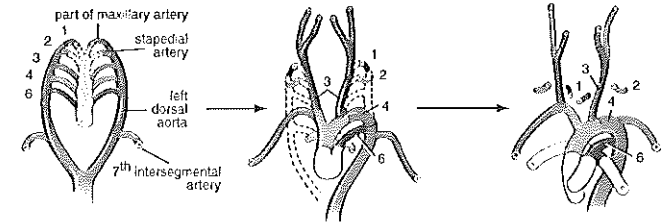


figure 12 Anterior vascular development showing the initially highly organized and symmetrical organization of the branchial arch arteries and dorsal aortae, which through numerous reorganizations (fusion, regression) transform into a complex non-symmetrical structure. Numbers and shades of gray indicate the corresponding pharyngeal arches. Modified from Larsen (1993).

Mesenchyme from pharyngeal arches 3, 4 and 6 is necessary for cardiac development, as described above, but the arteries running through the arches are needed for vascular development. Arch arteries are formed by endothelial cells, but neural crest cells are necessary for the separation of the primitive arteries from the endothelium. Crest cells ensheath the arteries, forming the smooth muscle cell layer, and are necessary for artery persistence, not formation (Waldo *et al.* 1996). The vascular system upstream of the heart is originally a bilaterally symmetrical system of aortic arches, but through extensive remodelling and degeneration this changes into a complex asymmetrical arterial network (figure 12).

Although in essence the same neural crest population as the cells forming the outflow tract, i.e. cardiac crest, it is possible to specifically inhibit the cells contributing to the aortic arches. Ablation of cardiac neural crest causes a spectrum of heart malformations in combination with aortic arch malformations (Kirby *et al.* 1985, Bockman *et al.* 1987, 1989, Kirby and Waldo 1990). Kirby *et al.* were able to separate the two events by antisense experiments, something not possible with ablations (Kirby *et al.* 1997). Back transplantation of crest cells treated *in vitro* with antisense oligos against specific *Hox* genes causes aortic arch malformations but no malformations in the outflow tract of the heart: the two events are independent.

11 4 MOLECULAR GENETICS OF NEURAL CREST

11 4 1 A COMPLEX CASCADING NETWORK OF GENES IS INVOLVED IN NEURAL CREST PATTERNING, DEVELOPMENT AND FATE DETERMINATION

Irrespective of whether neural crest cells get specified when still in the neural tube, during migration or at their final destination (and indications are that all three cases are applicable), in- and external cues determine the final fate. In the last few years a large number of genes have been implicated in the pattern formation of the head and thus in neural crest determination (table I). Some implications are merely based on the fact that the gene is expressed in either the neural crest cells themselves, their migratory pathway or their derivatives. Others genes have been more thoroughly investigated through the study of knock-out and transgenic mice, or through experiments affecting their expression or the activity of the proteins they code for. It is not feasible to give an exhaustive review of all the genes since the number of genes researcher think or know influence whichever part of neural crest cell development is mushrooming. This is possibly due to an increased interest in the ins and outs of neural crest cells, which in turn may be in recognition of the pivotal role of this group of cells in (head) development. Table I only gives a snapshot of the genes as known at the time of writing. I will concentrate on Homeobox proteins and the Endothelin system as these two classes have some bearing on the main theme of this thesis (CATCH22 and neural crest) and on ECM related components for reasons I will explain in that chapter. Before a detailed discussion of these subjects I will discuss some more general issues.

We have some knowledge about the molecules involved in the very first step of neural crest development, the time of induction. Epidermal ectoderm expresses Bone Morphogenetic Protein genes *BMP4* and *BMP7*, the products of which induce a dorsal fate (neural crest) and the expression of *BMP4* and *Dorsalin (dsl1)* in the neural plate, counteracting the ventralizing action of Sonic Hedge Hog (Shh) (Liem *et al.* 1995). A gradient of BMP (low in the midline, increasing laterally), regulated by the dorso-lateral mesoderm, is

neural crest of ganges. In the head region the neural tube is segmented normally. Gene expression in the neural tube, the layer in neuroepithelium and mesoderm segmentation is questionable. Evolutionary homology is thought about the head.

table I Genes implicated in (hindbrain) neural crest development owing to their expression pattern or the phenotype resulting from manipulation of their expression (knock-out, mutation). The REMARK column refers to a very broad category the gene belongs to in relation to paragraphs or sections of the introduction.

GENE	REMARK	REFERENCE
ADAM13	ecm	Alfandri <i>et al.</i> 1997
Ak3	homeobox	ten Berge <i>et al.</i> 1998
AP2		Morris-Kay 1996, Zhang <i>et al.</i> 1996, Schorle <i>et al.</i> 1996, Chazaud <i>et al.</i> 1996
barx1	homeobox	Tissier-Seta <i>et al.</i> 1995
BDNF	growth factor	Murphy <i>et al.</i> 1994
c-kit	ecm	Wehrle-Haller and Weston 1995
chn		Schilling <i>et al.</i> 1996
Chondroitin-6-sulphate	ecm	Oakley <i>et al.</i> 1994, Krull 1995
cKx2		Schutz and Niessing 1994
Collagen-IX	ecm	Ring <i>et al.</i> 1996
CRABP1	retinoid	Vaessen <i>et al.</i> 1990
Cx43	heart	Reaume <i>et al.</i> 1995, Ewart <i>et al.</i> 1997, Huang <i>et al.</i> 1998
dHAND, eHAND	heart	Srivastava 1997
Dlx2, 3 and 4		Robinson and Mahon 1994, Akimenko <i>et al.</i> 1994, Qui <i>et al.</i> 1995
EAP-300	heart	McCabe <i>et al.</i> 1995
Eck related		Brandt <i>et al.</i> 1995
En1		Augustine <i>et al.</i> 1995, Sadler <i>et al.</i> 1995
ErbB2 and 4	heart	Meyer and Birchmeier 1995, Lee <i>et al.</i> 1996, Glassman 1995
Evi1		Hoyt <i>et al.</i> 1997
Eya2		Duncan <i>et al.</i> 1997
FGF1 and 2	growth factor	Murphy <i>et al.</i> 1994, Zhang <i>et al.</i> 1997
GDNF	growth factor	Angrist <i>et al.</i> 1996
Gli2 and 3		Mo <i>et al.</i> 1997
Id2		Martinsen and Bronner-Fraser 1998
L10	heart	Kirby <i>et al.</i> 1995
LIF		Murphy <i>et al.</i> 1994
Mash1		Lo <i>et al.</i> 1994
matrix-metallo-proteases	ecm	Chin and Werb 1997
MFH-1		Iida <i>et al.</i> 1997
Ncx	homeobox	Hatano <i>et al.</i> 1997
Neuregulin	heart	Meyer and Birchmeier 1995, Lee <i>et al.</i> 1996, Glassman 1995
Neuronatin		Wijnholds <i>et al.</i> 1995
Nf1		Brannan <i>et al.</i> 1994
NGF	growth factor	Murphy <i>et al.</i> 1994
Notch 1, 2, and 3		Williams <i>et al.</i> 1995
NT3	growth factor	Murphy <i>et al.</i> 1994
Pax3	heart	Conway <i>et al.</i> 1997a, b
Pax7		Mansouri <i>et al.</i> 1996
PDGFalphaR	growth factor	Ho <i>et al.</i> 1994, Soriano 1997
PNA	ecm	Oakley <i>et al.</i> 1994, Krull 1995
PPEF		Montini <i>et al.</i> 1997
RA	retinoid	Ito and Montu 1995, Dupin and le Douarin 1995, Helms <i>et al.</i> 1997, Lee <i>et al.</i> 1995, Ozeki and Shirai (1998)
RAR	retinoid	Dickman <i>et al.</i> 1997
RET1		Schuchardt 1994, Robertson and Mason 1995
RXR	retinoid	Dickman <i>et al.</i> 1997
Scatter factor/hgf	growth factor	Takayama <i>et al.</i> 1996
Serotonin (5HT)	ecm	Moiseiwitsch and Lauder 1995, Choi <i>et al.</i> 1997
Serotonin receptor (SHT2B)	ecm	Moiseiwitsch and Lauder 1995, Choi <i>et al.</i> 1997

table I Continued.

GENE	REMARK	REFERENCE
Sey	ecm	Osumi-Yamashita et al. 1997
Snail 2		Thisse et al. 1995
Sox10		Southard-Smith et al. 1998
Steel	ecm	Wehrle-Haller and Weston 1995
tenascin-C	ecm	Tongiorgi et al. 1995
TGFbeta 1 and 2	growth factor	Zhang et al. 1997, Murphy et al. 1994, Sanford et al. 1997, Leblanc et al. 1995
Twist		Gitelman 1997, Stoetzel et al. 1995
urokinase type plasminogen activator	ecm	Agrawal and Brauer 1996
versican	ecm	Landolt et al. 1995
Xbap		Newman et al. 1997
XI-ets-1,-2 and XI-fl, qual Fli	ecm	Meyer et al. 1997, Mager et al. 1998
Zic3		Nakata et al. 1997

thought to be responsible for the induction of neural and neural crest fate in the ectoderm (Marchant et al. 1998). (Interestingly, BMP4 revisits the neural crest at a later stage when it induces a cell death program in rhombomere 3 and 5 crest (Graham et al. 1996, Marazzi et al. 1997)). Acting upstream of a cascade of genes involving members of the *Msx* family BMP sets up a pattern in the hindbrain (Sotokata and Maas 1994, Shimeld et al. 1996, Foerst-Potts and Sadler 1997, Houzelstein et al. 1997). Naturally there are more genes involved in this process, most notably those for the *Sek* family of receptor tyrosine kinases (Becker et al. 1994, Xu et al. 1995). Very early patterning of neural crest along the anterior-posterior axis (i.e. head-trunk division) is thought to be under influence of Hensen's Node. This process is likely to involve TGFbeta family proteins given the effect of TGF on *in vitro* cultured crest and node cells, the expression in crest cells and the effect of knock-out of the gene in mice (Murphy et al. 1994, Leblanc et al. 1995, Sanford et al. 1997). TGF is not the only growth factor involved in neural crest development. Given the wide range of derivatives of neural crest cells, it should not come as a surprise that growth factors play a role in their proliferation and fate choice processes. Table I lists growth factors involved in crest development. Retinoic Acid (RA) associated molecules are also listed in this table because it has been known for some time that RA has profound effects on craniofacial development. Overdose or depletion of RA during embryogenesis, or knock-out of its receptors (RAR and RXR) causes craniofacial deformities (Dickman et al. 1997), some not unlike CATCH22 type deformities. Neural crest related malformations in the heart are related to a different set of genes (table I).

Homeobox Genes

Homeobox or *Hox* genes are obvious candidates for the establishment of the identity of rhombomeres and their crest. Hox proteins contain a homeodomain, coded for by the homeobox (McGinnis et al. 1984, Scott and Weiner 1984) and thought to be involved in protein-DNA inter-

actions. Originally described in *Drosophila* (Lewis 1978), where they are involved in pattern formation, they were later found throughout the animal and plant kingdoms in the same capacity. They act as transcription factors by binding to DNA. *Antennapedia* class *Hox* genes are arranged in clusters in the genome, with an evolutionary conserved linear arrangement. Higher organisms (human, mouse) have four clusters, thought to have arisen from one cluster by duplication events. Primitive organisms have one, two or three clusters, roughly corresponding to their position in the evolutionary tree (Pendleton et al. 1993). Paralogs (of which there can be up to four in higher organisms) share a high degree of homology, which has implications for compensatory effects when one of them is not functional (i.e. knock-out or mutant mice). The linear order of the genes in a given paralogous group reflects the anterior expression boundaries of the constituent genes. The more 3' the gene, the more rostral the anterior expression boundary. Posteriorly, expression often extends far caudally. As an effect, caudally the expression domains of many *Hox* genes overlap, but anteriorly segments have unique *Hox* codes (figure 5 shows the expression domains of some *Hox* genes in relation to rhombomeres). Some *Hox* genes are not expressed caudally and are restricted to the hindbrain (Godsave 1994). Anterior boundaries coincide with rhombomere boundaries, so it is reasonable to assume a role for hox genes in setting up or maintaining rhombomere identity. And indeed, manipulation of *Hox* codes often results in hindbrain and neural crest maldevelopment (Rijli et al. 1998). A knock-out mouse defective in *Rae28* (a homolog of the *Drosophila polyhomeotic* gene) suffers from parathyroid and thymic hypoplasia, cardiac abnormalities, cleft palate and more. In embryos the anterior boundaries of several *Hox* genes, including *hoxb3* and *b4* in the hindbrain region, had shifted rostrally (Takahara et al. 1997). Aortic arch malformations can also be induced by treating isolated crest cells with antisense oligos against specific *Hox* genes, and back transplanting these cells isotopically (Kirby et al. 1997).

An interesting question is whether the *Hox* gene expression pattern in the hindbrain can be influenced by transplantations. Is the



expression fixed or will the pattern be remodeled by the new environment the cells find themselves in? Saldívar *et al.*, amongst others, sought to answer this question by rotating the rhombomere 4+5 section so that anterior and posterior were reversed (Saldívar *et al.* 1996). *Hoxa3* normally starts at the boundary between rhombomere 4 and 5 but rhombomere 5 crest cells joining the stream of rhombomere 4 do not express the gene. The caudally migrating rhombomere 5 cells that join the rhombomere 6 stream do, however. When, as the result of the transposition, rhombomere 5 is where rhombomere 4 used to be, neural crest cells from rhombomere 5 migrate to the proper rhombomere 4 targets. Rhombomere 4 crest cells that would normally have populated the targets do not express *Hoxa3*, but the ectopic rhombomere 5 cells now populating the targets express the gene in the ganglia, although not in branchial arch 2. Placing the more caudally located *Hoxa3* positive rhombomere 6 in the rhombomere 4 position results in *Hoxa3* expression in arch 2, its ganglia and the migrating crest cells. So at the stage the procedures were performed, the *Hox* code in rhombomere 6 is cell autonomous, while in rhombomere 5 it is partially influenced by the ectopic surroundings. Prince and Lumsden (1994) and Couly *et al.* (1998) observed cell-autonomous expression in their transplantation experiments as well.

Hox genes obviously do not work alone, but are part of a network of genes (gene products). For example, *Hoxa3* and *Pax1* appear to cooperate in thymic development. *Hoxa3* and *Pax1* knock-out mice have some phenotypic overlap: the former are athymic and thyroid hypoplastic, while the latter have an underdeveloped thymus. *Pax1* knock-outs of course lack expression of *Pax1* in the prospective thymus, but *Hoxa3* knock-outs are also affected with poor expression of *Pax1* in what would have become the thymus. Migration and proliferation of crest cells appear normal, so presumably *Hoxa3* directly or indirectly maintains *Pax1* expression for proper crest cell differentiation and thymus development (Manley and Capocchi 1995). This link between Pax and Hox is not the only occasion of two components that are involved in neural crest development individually and turn out to be active in the same pathway or network.

structure (Bosma, cartilage), but also to part of the heart and its outflow tract. Neural crest cells, migrating from the rhombomeres, yield most of the mesenchyme of these arches. Secondary vesicles like the thyroid, parathyroids and thymus originate from the pharyngeal pouches (Bosma, floor of mouth).

Earlier, I briefly mentioned RA as a molecule that can seriously affect craniofacial development. Hox and RA pathways seem to mesh at the level of *Hoxb1* expression regulation. Normally restricted to rhombomere 4 by a repressor which gene contains a RARE (RA Response Element), *Hoxb1* expression spreads into neighboring rhombomeres when the RARE of the repressor gene is mutated (Studer *et al.* 1994). Another homeobox gene providing evidence for a connection between the RA and homeobox pathways is *Gooseoid*. *Gooseoid* has a bi-phasic expression pattern: very early in embryogenesis in the Spemann organizer, later in, amongst other sites, some neural crest tissues (the first and second branchial arches, the frontonasal mass) (Gaunt *et al.* 1993). Mice with homozygous *Gooseoid* null mutations do not exhibit any problems related to the first phase of expression but do show a second phase related phenotype during organogenesis. The nasal and auditory system and the lower mandible and its musculature are affected and the mice die perinatally (Rivera-Perez *et al.* 1995, Yamada *et al.* 1995, 1997). RA treated embryos have auditory phenotypes resembling those in the *Gooseoid* knock-out mice, which prompted Zhu *et al.* to examine the *Gooseoid* expression pattern in RA treated embryos. They did indeed find a change in its expression in the branchial arch region, providing us with another example of the joining of RA and homeobox pathways in craniofacial neural crest regulation (Zhu *et al.* 1997). The experiments I described above place RA upstream of *Hox* genes. Homeobox pathways also network with each other: for example, ectopic *Hoxa1* expression transforms the second rhombomere into a fourth rhombomere and induces ectopic expression of *Hoxb1* in this transformed rhombomere (Zhang *et al.* 1994).

Homeobox genes are not limited to the established *Hox* clusters. Many isolated homeobox genes have been isolated, and knock-outs of several of these show abnormal development of (hind-brain) neural crest derivatives. For example, *Mhox* knock-out mice have multiple skeletal deformities, including deformities of elements derived from neural crest (Martin *et al.* 1995). Together with the expression pattern in the chicken (Kuratani *et al.* 1994) this suggests that *Mhox* is involved in chondrogenesis, possibly in epithelio-mesenchymal interactions. See table I for more examples.

// # i 2 **Endothelin System**

Endothelins are small 21 amino acid proteins, of which there are three types (EDN1, 2 and 3, also called ET-1, etc.) (Inoue *et al.* 1989). They are ligands for Endothelin Receptors EDNRA and EDNRB (or ET-A and ET-B) which belong to the G protein-coupled receptor family (Arai *et al.* 1990, Sakurai *et al.* 1990). Receptor A binds ET-1 and ET-2, while receptor B binds all three Endothelins. The third component of the Endothelin system consists of the Endothelin Converting Enzymes ECE1 and ECE2 (Xu *et al.* 1994, Emoto and Yanagisawa 1995). These membrane bound metalloproteases convert so-called big Endothelin into mature Endothelin by a highly specific cleavage reaction.

(Reedy *et al.* 1998). Neural crest migration patterns can also be influenced by influencing the fate of neural crest cells by over-expression of N-myc (Wakamatsu 1997). Erickson and Reedy call this differentiation state \leftrightarrow pathway interaction the phenotype-directed model of neural crest cell migration (Erickson and Reedy 1998).

Immunohistochemistry and *in situ* histochemistry point to many ECM associated molecules with a differential expression pattern coinciding with the crest cell migration pattern, or with an influence on migration after perturbation of their normal expression (Boucaut *et al.* 1984, Bronner-Fraser 1985, Bronner-Fraser 1986, Krotoski *et al.* 1986, Poole and Thiery 1986, Tan *et al.* 1987, Bronner-Fraser 1988a, Bronner-Fraser and Lallier 1988, Stern *et al.* 1989, Poelmann *et al.* 1994). Receptor-ligand complexes like the ephrin receptor-ephrin system play a role in anterior-posterior regionalization of the somite and in the appropriate response of crest and neuronal cells. *In vitro* tests in chicken show that ephrin receptors in motor axons and neural crest cells are vital in recognizing environmental signals (ephrin ligands in the caudal somite) and eliciting the appropriate response (Krull *et al.* 1997, Wang and Anderson 1997). In *Xenopus* the Eph family was implicated in cranial neural crest cell migration because of differential expression of ephrin family member *Pagliaccio* (Winning and Sargent 1994). Smith *et al.* were able to manipulate migratory behavior of *Xenopus* cranial crest cells by perturbing the Eph system, showing it is responsible for keeping the streams of hindbrain neural crest cells separate (Smith *et al.* 1997). We can easily derive that the ephrin ligand-receptor system is involved in some aspect of cell migration, probably the mediation of the collapse of active membrane protrusions. Actually, the implication of the ephrin system in avoidance behavior of neural crest cells is not surprising given that the Eph system functions in neurogenesis by mediating growth cone repulsion and neural crest cells are migratory cells of neural descent.

Various extra-cellular matrix components and their receptors, as well as molecules involved in cell-cell contact have been shown to regulate crest cell migration (see also **table I**). Integrins, fibronectin,

laminin and cadherins are among the most important regulators of selective neural crest cell migration. Integrins interact with laminin, fibronectin and cadherin, with differential affinities for different subfragments and isoforms and cellular morphology dependent on the specific integrin-ligand pair (Perris *et al.* 1996, Beauvais-Jouneau *et al.* 1997). Migratory behavior changes in time, concurrent with the change of distribution of at least one of the integrins (Desban and Duband 1997). More proof of the importance of integrin is supplied by the craniofacial deformities displayed integrin knock out mice (Goh *et al.* 1997) and neural crest defects induced by antisense oligonucleotide treatment (Kil *et al.* 1996). These results underline the importance of integrins for proper neural crest cell migration.

The transition from stationary to migratory cells (Epithelial-Mesenchymal Transition, EMT) involves changes in cell-cell contact (decrease) and cell-ECM contact (elevation). Re-establishment of cell-cell contact at transition of migratory to stationary status upon arrival in for example a ganglion, involves more or less the reverse changes. This is the domain of cadherins and catenins. Differential expression (dependent on fate or destination) of various Cadherin family members in crest cells undergoing epithelial-mesenchymal transition shows the involvement of the protein family in crest development (Nakagawa and Takeichi 1995, Simonneau *et al.* 1995, Inoue *et al.* 1997). Over-expression of cadherins in the neural tube has serious negative consequences for neural crest cell migration (Nakagawa and Takeichi 1998). Integrins are one of the first steps in a signalling pathway involving integrins, homeobox proteins, slug, kinases, phosphatases, fibronectin, catenin and cadherin (LeDouarin *et al.* 1994, Monier-Gavalle and Duband 1995, Newgreen and Minichiello 1995, Inoue *et al.* 1997, Monier-Gavalle and Duband 1997). A coordinated sequence of events involving these molecules balances migration and anti-migratory cell-cell contact.

From above it is clear that the ECM has a significant influence on neural crest behavior, but the interaction between crest cells and the ECM goes two ways: *in vitro*, crest cells modify the laminin substrate on which they migrate (Lallier *et al.* 1994). We have also seen that repulsion plays an important role in guiding crest cell migration in the trunk. But Stern *et al.* (1991a) have shown, by rotating the neural tube dorso-ventrally, that the dorsal root ganglia are still generated at their normal position in the anterior sclerotome when the neural crest cells emanate from the ventral aspect of the neural tube, pointing to positive (attractive) signals. Positive and negative cues guide the neural crest cells on the path towards their final destination.

In conclusion, predetermination influences pathway choice (phenotype-directed neural crest cell migration) but, conversely, determination is influenced by the pathway choice, which in turn is influenced by the environment (environment-directed model of neural crest morphogenesis).

11 5 DEPENDING ON THE AXIAL LEVEL, EXTIRPATED NEURAL CREST IS COMPENSATED FOR BY REMAINING NEURAL TISSUE

Regulation of migration, determination and differentiation of (cranial) crest cells is a very complex system involving intricate networks of interacting gene products. However sensitive such a system might appear, the cephalic neural crest is surprisingly flexible. Extirpation of sizeable segments of neural crest, uni- or bilaterally, induces extensive compensatory reactions from remaining neural crest and neural tube. This compensatory behavior is region and time specific. When the rhombencephalic neural crest is ablated, branchial arches still receive their neural crest contribution and morphology seems to be normal after 48 hours. The *Hox* code (at least for those genes tested) particular to the branchial arches is intact (Hunt *et al.* 1995). Re-appositioning of neuroectoderm and surface ectoderm is essential for regeneration. This contact induces production of Pax3 and, later, slug (Sechrist *et al.* 1995, Buxton *et al.* 1997). The regenerative powers of the caudal midbrain and hindbrain are stronger than caudal forebrain and rostral midbrain. Possibly time is a factor here: at earlier stages the difference seems to be smaller. Not surprisingly, bilaterally ablated embryos have more difficulties compensating for the lost tissue than unilaterally ablated embryos. In both cases the structures receiving neural crest contributions seem to be reduced in size (Sechrist *et al.* 1995). Crest is regenerated from rostral, caudal and, when available, lateral neural crest. Rostral and caudal cells migrate longitudinally and then into branchial arches, without changing their native *Hox* code (Couly *et al.* 1996). Other ablation experiments do show a change in *Hox* gene expression, however. Ablation at the 10-12 somite stage of dorsal rhombomeres 5 + 6 causes redirection of some rhombomere 4 and 7 crest cells into branchial arch 3 (whereas these cells normally migrate to arch 2 and 4 respectively) (Saldivar *et al.* 1997). There is an upregulation of *Hoxa3* in the redirected rhombomere 4 cells, and *Krox20* is upregulated in caudal rhombomere 3, rhombomere 4 and rostral rhombomere 7. Hyoid (partially derived from arch 3, which receives crest cells from posterior rhombomere 5 and from rhombomere 6) development is normal. So, ablation seems to cause proper re-specification of neural crest cells. In contrast, when the neural tube containing rhombomeres 4 through 6 is rotated, crest cells migrate to their new branchial arch without for example up regulating *Hoxa3* for arch 3, which results in hyoid malformation. Hunt *et al.* did a detailed study of the effects of rotation of rhombomere segments on *Hox* gene expression, crest cell migration and branchial arch development and conclude that it is highly dependent on which segment is rotated (Hunt *et al.* 1998). An r3-7 rotation would result in stable (cell-intrinsic) expression of *Hox* genes and hence branchial arch malformation, whereas an r1-7 rotation would result in re specification of the code and generally normal arch derivatives.

In contrast to the strong regenerative powers of cephalic neural crest up to the otocyst, the response of postotic neural crest to ablation is much less robust. After somite 6 stage no regenera-

tion of cardiac neural crest from anterior or posterior crest is observed, nor compensatory migration from remaining neural tube. Before that stage only very little compensation occurs. In the trunk the picture is the same, although differentiated crest cells in the form of melanocytes do migrate to fill in the gaps (Suzuki and Kirby 1997).

III DISEASES AND SYNDROMES

III 1 SYNDROMES RELATED TO HINDBRAIN NEURAL CREST MALDEVELOPMENT

From the introduction on neural crest development it is clear that neural crest contributes to a wide variety of structures throughout the vertebrate body. This is mirrored in the broad spectrum of diseases where (part of) the phenotype can be traced back to failure in neural crest development. I will highlight just two (related) examples: Hirschsprung's Disease and Waardenburg Syndrome.

Hirschsprung's Disease (HSCR) is characterized by the failure of neural crest cells to develop ganglia in longer or shorter segments of the digestive tract (Peters-van der Sanden 1994 (thesis)). The consequent absence of peristaltic movement and hence accumulation of feces is often fatal. Enteric ganglia are formed by hindbrain neural crest cells, as described earlier. In recent years no less than seven genes have been implicated by virtue of being deleted or mutated in HSCR patients or mice with an HSCR phenotype. They are the genes for Glial Derived Neurotrophic Factor (*GDNF*) (Ivanchuk *et al.* 1996, Moore *et al.* 1996, Pichel *et al.* 1996, Sanchez *et al.* 1996), Endothelin Receptor B (*EDNRB*) (Kusafuka *et al.* 1996), *RET* (Luo *et al.* 1993, Ederly *et al.* 1994, Romeo *et al.* 1994), *SOX10* (Southard-Smith *et al.* 1998), Endothelin Converting Enzyme 1 (*ECE-1*) (Yanagisawa *et al.* 1998), Neurturin (Doray *et al.* 1998) and Endothelin 3 (*ET-3*, *EDN3*) (Bidaud *et al.* 1997, Kusafuka *et al.* 1997, Robertson *et al.* 1997). Six of these genes are physiologically closely linked: Endothelin 3 is a ligand for *EDNRB* and a substrate for *ECE-1*, Neurturin is a ligand for *RET*, and *GDNF* is a ligand for the *RET/GDNFR*-alpha receptor complex (Ivanchuk *et al.* 1996). Interestingly, *RET* mutations are also linked to other neuro-cristopathies: one of the many neural crest related cancers, Familial Medullary Thyroid Carcinoma (FMTC), is caused by specific mutations in the *RET* gene (Donis-Keller *et al.* 1993). Other mutations in the *RET* gene cause MEN2A and MEN2B (Multiple Endocrine Neoplasia type 2), two other types of crest derived cancers (Donis-

Table 2 Clinical syndromes discussed in the context of single and multigene aneuploidy syndromes and a short description of their respective phenotypes.

SYNDROME	PHENOTYPE
Alagille	intrahepatic bile duct paucity, cardiac, ocular, skeletal and facial defects
Angelman	ataxic gait; hypotonia; neurological disorders (epilepsy, seizures, motor and mental retardation, absence of speech, abnormal EEG); facial deformities
CHARGE	Coloboma; Heart disease; Atresia of choanae; mental Retardation; Genital hypoplasia; Ear abnormalities
Goldenhar	malformations of the vertebrae (e.g. fused vertebrae) and face and cranium (microcephaly, hydrocephalus, skull defects, microtia and other ear defects, eye problems, uni- or bilateral facial hypoplasia, mandibular abnormalities); often combined with heart defects
Greig	craniofacial abnormalities; poly- and/or syndactyly of hands and feet
Hirschsprung	intestinal aganglionosis
Holoprosencephaly	failure of midline forebrain structure cleavage, causing cyclopia or hypotelorism and other craniofacial deformities
LGS/TRPS II	sparse scalp hair; bushy eyebrows; large protruding ears; broad nasal bridge and bulbous nose; elongated upper lip; cone-shaped epiphyses; mental retardation; multiple cartilaginous exostoses (outgrowths or protuberances) and frequently other facial deformities
Rubinstein-Taybi Syndrome (RTS)	facial abnormalities; broad big toes and thumbs; neurological defects (mental, motor, speech and social retardation)
Waardenburg Syndrome type 1 (WS1)	patchy depigmentation (e.g. white forelock); dystopia canthorum; mild facial dysmorphism; auditory problems
Williams	facial and cardiovascular abnormalities; growth retardation; neurological anomalies (cognitive development delay)
Xp22 deletion	short stature; X-linked recessive chondrodysplasia punctata; X-linked ichthyosis; mental retardation; Kallman syndrome (= eunuch like physiology due to lack of gonadotropin releasing hormone)

Keller *et al.* 1993, Mulligan *et al.* 1993, Hofstra *et al.* 1994, van Heyningen 1994). Many other tumors of neural crest derivatives have been discovered, for example neurofibromatosis types 1 and 2 (NF1, NF2), neuroblastoma, and pheochromocytoma.

A syndrome that sometimes combines with HSCR and where various neural crest derivatives can be affected, is the Waardenburg Syndrome. Four different types are recognized, imaginatively referred to as WS1, WS2, WS3 (Klein-Waardenburg) and WS4 (Shah-Waardenburg). The phenotype for *PAX3* associated WS1 is described in table 2. WS2 is very similar to type 1 but without the dystopia of the inner canthi and with premature graying rather than white forelock. Type 3 is characterized by hypoplasia of limb muscles and contractures of elbows and fingers, in combination

genes is very wide, ranging from limb malformations and neurological disorders to heart and skin anomalies. Two very common elements not described by the acronym are cleft palate and ear/hearing problems. The secondary palate forms from the same population of neural crest cells that forms the nose.

Table 3 Single gene syndromes or disorders, their genes and the chromosomal map position of their genes.

SYNDROME	GENE	MAP POSITION	NOTE	REFERENCE
Alagille Syndrome	Jagged1 (JAG1)	20p12		Oda <i>et al.</i> 1997 Krantz <i>et al.</i> 1997a, b
Angelman Syndrome	E6-AP ubiquitin protein ligase (UBE3A)	15q11-q13	*	Meijers-Heijboer <i>et al.</i> 1992 Budarf and Emanuel 1997 Kishino <i>et al.</i> 1997 Matsura <i>et al.</i> 1997
Greig Syndrome	GLI3	7p13		Wild <i>et al.</i> 1997
Holoprosencephaly	Sonic Hedgehog (SHH)	7q36	**	Roessler <i>et al.</i> 1997a, b
Rubinstein-Taybi Syndrome (RTS)	CREB Binding Protein (CBP)	16p13.3		Breuning <i>et al.</i> 1993 Hennekam <i>et al.</i> 1993 Masuno <i>et al.</i> 1994 Petrij <i>et al.</i> 1995 McGaughan <i>et al.</i> 1996 Akimaru <i>et al.</i> 1997 Wallerstein <i>et al.</i> 1997
Waardenburg Syndrome type 1 (WS1)	PAX3	2q35-q37		Lu-Kuo <i>et al.</i> 1993 Yassabehji <i>et al.</i> 1993, Baldwin <i>et al.</i> 1995 Read <i>et al.</i> 1997
Aniridia and other eye disorders	PAX6	11p13	***	Francke <i>et al.</i> 1979 Glaser <i>et al.</i> 1992 Davis <i>et al.</i> 1993 Ekker <i>et al.</i> 1995 Fantes <i>et al.</i> 1995 MacDonald <i>et al.</i> 1995 Schedt <i>et al.</i> 1996 Axton <i>et al.</i> 1997

* Besides deletion and mutation also improper imprinting

** Chromosomal rearrangements not directly affecting the structure of the Shh gene still have influence on its regulation

*** Also aneuploidy in form of polyzygosity

with varying type 1 features. Lastly, type 4 is the Waardenburg syndrome characterized by the occurrence of Hirschsprung's disease in combination with the auditory-pigmentary features (though association of WS2 with HSCR has been reported as well (van Camp *et al.* 1995)). Endothelin 3 and its receptor EDNRB are what links WS4 to HSCR (see Read and Newton 1997 for a review on Waardenburg Syndrome).

A part of the nose and the philtrum. Facial development, in higher vertebrates, the palpe commissure, namely of secondary palates, involves the raising and falling of two palatine shelves, and this process is influenced by the upper lip (mandible), which one also of neural crest origin. So middevelopment of various

Other syndromes, some of which are discussed in **chapter III-2**, have phenotypes that combine craniofacial malformations with heart defects (for example CHARGE and Goldenhar Syndrome) (Bolande 1974, 1981, 1984, Couly 1981). Like the syndromes discussed here these syndromes have a pleiotropic and variable phenotype affecting many different neural crest derivatives, but also apparently unrelated structures. More about heart development from a neural crest perspective in **chapter III-3-2-1**. Congenital ear abnormalities are also a recurring feature of syndromes (for example Townes-Brocks, Wildervanck, Smith-Magenis, and Turner syndromes and aforementioned Goldenhar, CHARGE and Waardenburg) (Gorlin *et al.* 1995). Not so surprising maybe, considering the fact that ear development is quite a complex process involving neural crest, mesoderm and ectoderm. Where genes have been isolated, these syndromes can provide us with important clues to the complex question of neural crest regulation, even if only for a subset of crest cells (like those contributing to the heart or ears).

III 2 SINGLE GENE AND CONTIGUOUS GENE HAPLOINSUFFICIENCY OR AUTOSOMAL SEGMENTAL ANEUSOMY SYNDROMES

When dealing with the fairly complex and heterogeneous neural crest system, soon the question arises whether crest related anomalies, especially the variable and complex ones, are single or multigene defects. "Simple" phenotypes such as carcinomas are usually linked to one gene, but as described above with HSCR, the same cancer causing gene can be involved in a syndrome as well. Some forms of HSCR are single gene haploinsufficiency syndromes. Single gene syndrome in the sense of one gene being necessary and sufficient for causing the syndrome; different genes, on their own, can be causative for the same phenotype as is the case with HSCR.

There are a number of syndromes for which we can say, with varying degrees of certainty, that they are single gene haploinsufficiency deletion syndromes (**table 3**) (Fisher and Scambler

1994, Engelkamp and van Heyningen 1996, Budarf and Emanuel 1997). Like HSCR, many of these have a heterogeneous genetic etiology in that the syndrome is caused by either physical or functional deletion (functional deletion = mutation or regulation defect) of a single gene. The majority of these syndromes have complex cardiac and/or craniofacial abnormalities that could involve neural crest directly or indirectly (RTS, WSI, Allagille, Angelman, Greig). These syndromes show that complex syndromes are not necessarily caused by complex genotypes.

In contrast to the relatively large number of single gene haploinsufficiency syndromes, not many contiguous gene haploinsufficiency syndromes have been described. Contiguous gene syndrome is a description coined by Schmickel (1986) to describe disorders resulting from the deletion or duplication of adjacent genes in a given chromosomal region. Phenotypes linked to each individual gene would exist, and the syndrome phenotype would be the sum of these phenotypes. Consequently, the variability of the phenotype would correlate with the extent and position of the deletion/duplication. I will discuss two syndromes with a craniofacial component of which we are reasonable sure that they are multigene syndromes.

The first example is the Langer-Giedion Syndrome or Tricho-Rhino-Phalangeal Syndrome type II (LGS/TRPS II) (table 2). The 8q24 chromosomal region associated with the syndrome harbors at least two genes (*TRPS* and *EXT1*, respectively) that contribute to the associated partial syndromes TRPS I and hereditary multiple exostoses (Ahn et al. 1995). Since neither of these two types of patients display the mental retardation component of LGS/TRPS II, a third gene is implicated (Ludecke et al. 1995). LGS/TRPS II has all the signs of indeed being a contiguous gene syndrome.

The second example is Williams Syndrome (WS), a disorder that includes facial and cardiovascular abnormalities associated with deletions in 7q11.23. Seven transcription units have been mapped to the deletion (Frangiskakis et al. 1996, Osborne et al. 1996, Peoples et al. 1996, Budarf and Emanuel 1997) and families with a partial phenotype — cognitive disorder and supravalvular aortic stenosis (SVAS) — have been discovered to have a smaller deletion

encompassing only two of these genes: *Elastin* (*ELN*) and *LIM-Kinase I* (*LIMK1*). Elastin has been linked to SVAS (Ewart et al. 1993b) and LIM-Kinase I to the cognitive disorder (Frangiskakis et al. 1996). It is therefore reasonable to assume that hemizygoty of *ELN* and *LIMK1* plus one or more additional genes is causative for WS and thus that WS is a contiguous gene syndrome.

For a final example I will digress a little from the haploinsufficient contiguous gene syndromes to a recessive contiguous gene syndrome because it is a textbook example of a contiguous gene syndrome. The case in point is the Xp22 deletion syndrome. Male patients show a phenotype when nullisomic for the region, their phenotype depending on the length of the deletion (Ballabio et al. 1989). Up to five different diseases can be associated with the syndrome (table 2). Molecular analysis of various deletions showed that the aggregate phenotype is the sum of the phenotypes caused by the deleted loci. Many different combinations of diseases are found since these deletions can be either interstitial or terminal. Three of the diseases have a gene assigned: ichthyosis (*Arylsulfatase A*), chondrodysplasia punctata (*Arylsulfatase E*) and Kallman syndrome (*Kalig 1*) (Ballabio and Andria 1992, Winter 1996). Mutations in these genes cause the isolated forms of the respective diseases (Winter 1996).

Deletion syndromes such as Turner, Prader-Willi, van der Woude, Zellweger, Goltz, Smith-Magenis, and Wolf-Hirschhorn cannot yet be assigned to the single gene or contiguous gene disorder classes (Fisher and Scambler 1994, Budarf and Emanuel 1997).

III 3 CHROMOSOME 22Q11 DELETION SYNDROME: NEURAL CREST AND CHROMOSOME 22

Craniofacial and cardiac maldevelopment can be a component of both single and multigene defects, as shown above, so these cannot tell us anything about the genetic nature of a syndrome that has cardiac and craniofacial components. However, if one were to choose as a subject such a syndrome associated with large chromosomal deletions, one could expect to hit upon a contiguous gene syndrome. CATCH22 or chromosome 22q11 Deletion Syndrome is an excellent candidate for a haploinsufficient contiguous gene syndrome because it marries a variable phenotype to a variable genotype. By studying the deletions associated with the syndromes and the genes located in the deletions, we are likely to find at least one but possibly more genes involved in some part of the development of the neural crest. With Williams syndrome, which could have some neural crest connection since it has facial and cardiovascular abnormalities, it has been possible to assign individual genes to at least two parts of the phenotype, which shows the value of studying the genetics in relation to the clinical aspects of a syndrome. CATCH22 genetics is the central theme of this thesis, so that aspect will be discussed in detail in the results section. The clinical aspects will be described here, after the following brief introduction on some chromosome 22 specifics.

III 3 1 VARIOUS DISORDERS ARE ASSOCIATED WITH CHROMOSOME 22 REARRANGEMENTS

Chromosome 22 is the smallest of the human autosomal chromosomes. The total size is approximated at less than 56 Mb (figure 13) and sequencing was completed in 1999 (Dunham *et al.* 1999). Chromosome 22 contains several important regions of chromosomal rearrangements associated with clinical syndromes (Dumanski *et al.* 1991). For example: the so-called Philadelphia chromosome, a reciprocal t(9;22) translocation resulting in fusion of the BCR (chromosome 22) and ABL (chromosome 9) genes, is associated with multiple forms of leukemia (ALL (Acute Lymphoblastic Leukemia), CML (Chronic Myeloid Leukemia) and AML (Acute Myelogenous Leukemia)) (reviews in Heisterkamp and Groffen 1991, Mitelman 1993, Glassman 1995, O'Brien *et al.* 1997a, Saglio *et al.* 1997). Burkitt's Lymphoma, another hematologic malignancy, can be related to t(8;22) translocations that bring chromosome 8 *c-Myc* under the control of chromosome 22 *Immunoglobulin Lambda Light Chain* regulatory regions (Rowley 1982, Aisenberg 1984, Emanuel *et al.* 1986, Zeidler *et al.* 1994). Predisposition to breast cancer may be linked to the constitutional t(11;22)(q23;q11) translocation (Lindblom *et al.* 1994), while a slightly different t(11;22)(q23;q11) is associated with Ewing sarcoma and peripheral neuroepithelioma (van Kessel *et al.* 1985, Griffin *et al.* 1986, Budarf *et al.* 1989). A different kind of rearrangement, resulting in the duplication of 22pter-q11, causes CES (Cat Eye Syndrome) (Schinzel *et al.* 1981, Mears *et al.* 1994). Chromosome 22 further contains proto-oncogenes, tumor suppressor genes and other genes potentially involved in neoplasia. Amongst them are *LIF1* (Sutherland *et al.* 1989), *c-Cis* (*PDGF-2(B)*) (Antoniades 1991), *NF2* (Rouleau *et al.* 1990, Arai *et al.* 1994), *GAR22* and *RRP22* (Zucman-Rossi *et al.* 1996), and *p300* (Muraoka *et al.* 1996). For an overview of the many other genes located on chromosome 22, from *ATP-dependent DNA Helicase II* to *X-box binding protein I*, see URL1.

III 3 2 CHROMOSOME 22Q11 ANOMALIES ARE ASSOCIATED WITH A VARIETY OF CARDIOVASCULAR AND CRANIOFACIAL ABNORMALITIES, MANY OF WHICH ARE CRANIAL NEURAL CREST RELATED

Relevant to this thesis, the most important anomalies on chromosome 22 are in the q11 region. These have our interest because they are found in patients that suffer from the various syndromes that together form the CATCH22 syndrome. CATCH22 is an acronym for Cardiac abnormalities, Abnormal face, Thymic hypoplasia or aplasia, Cleft

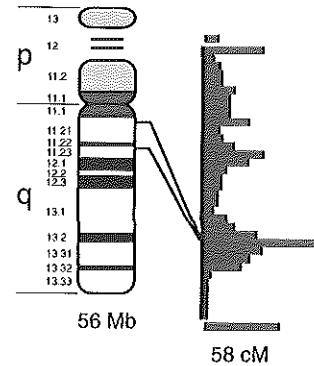
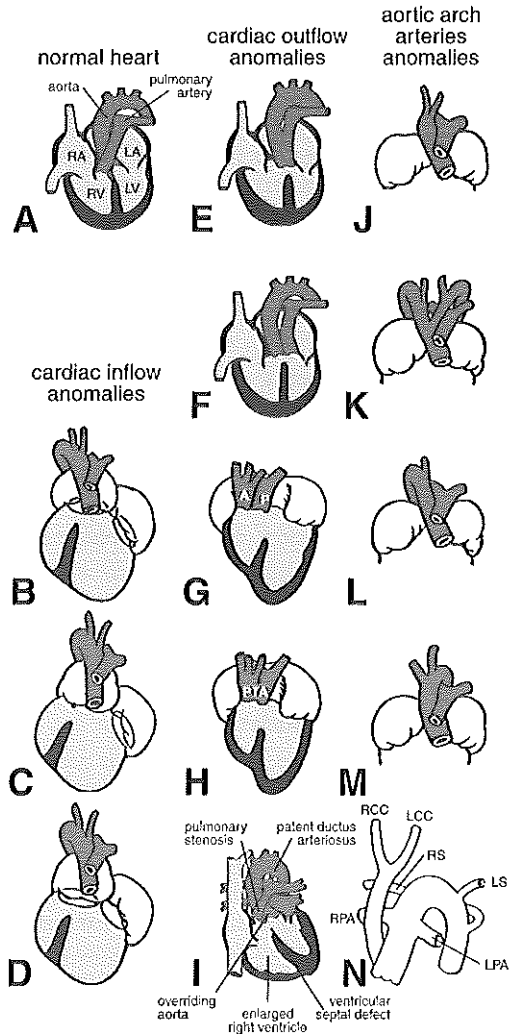


figure 13 Schematic depiction of a metaphase human chromosome 22. The short p arm and the long q arm are shown, as well as the various subbands and the pericentromeric region (dark gray). A bar graph on the right depicts relative gene density on a 58 centimorgan map. After URL1.

palate and Hypocalcemia (secondary to hypo-/aplasia of parathyroids) and 22q11 deletion. This acronym (taken from the title of Joseph Heller's book (Heller 1962)) describes typical abnormalities found in patients (although the real spectrum is much wider and continues to widen). CATCH22 is actually a group of overlapping syndromes (Wilson *et al.* 1993): DiGeorge (DGS, OMIM #188400) (de la Chapelle *et al.* 1981, Carey *et al.* 1990, Driscoll *et al.* 1992a), Velo Cardio Facial (VCFS, OMIM #192430)/Shprintzen (Driscoll *et al.* 1992b, Scambler *et al.* 1992, Burn *et al.* 1993, Kelly *et al.* 1993), Conotruncal Anomaly Face (CTAF)/Takao (Kinouchi *et al.* 1976, Takao *et al.* 1980, Shimizu *et al.* 1984), and Opitz-G (OMIM #300000) (Robin *et al.* 1995, McDonald-McGinn *et al.* 1995). All these syndromes have certain phenotypical features in common and are associated with a number of chromosomal rearrangements in the 22q11 region. CATCH22 is also referred to as 22q11 Deletion Syndrome (22DS) because some patients and patients' families object against the connotations associated with the CATCH22 acronym ("A Catch 22 situation is one where you cannot do one thing until you do another thing, but you cannot do the second thing until you do the first thing. It is therefore impossible for you to do anything" Collins COBUILD English Language Dictionary (1990)). In either case the name reflects one of the two central features that binds the constituent syndromes together: deletion of chromosome 22(q11). Hemizyosity caused by these deletions is somehow responsible for the quite diverse spectrum of phenotypes described for the constituent clinical syndromes.

DiGeorge Syndrome (DGS), which may be considered to lie at the severe end of the spectrum of CATCH22, is characterized by hypoplasia/aplasia of the thymus and parathyroid glands, conotruncal heart defects and varying craniofacial dysmorphologies (Conley *et al.* 1979). The syndrome is named after A.M. DiGeorge, who first described the link between absence of the thymus and cellular immunodeficiency syndrome in these patients (DiGeorge 1968). DGS can express itself in many grades of severity. Complete DiGeorge syndrome is defined as thymus aplasia; in cases of thymus hypoplasia we should, strictly speaking, talk about partial DGS



The heart shown as two simple tubes that flow, but through compartmentalization, folding, and rotation, a complete four-chambered heart is formed. Heart defects are commonly found in CATCH22 patients, or nonsyndromic patients with 22q11 deletion. Structurally, the heart does combine very little.

(Lischner 1972). Velo-Cardio-Facial Syndrome (VCFS) or Shprintzen syndrome overlaps with DGS in the cardio-facial part but adds neurological disorders (learning disabilities, behavior difficulty, mild mental retardation, psychosis) and cleft palate to its phenotype (Shprintzen *et al.* 1978, Arvystas and Shprintzen 1984, Goldberg *et al.* 1985, Driscoll *et al.* 1992b, Goldberg *et al.* 1993). Some consider DGS and VCFS to be the same syndrome with different severity, or VCFS the familial form of DGS. Most DGS cases are sporadic, but familial cases have been reported, such as a DGS child of a VCFS parent (Stevens *et al.* 1990). In non-sporadic cases there have been reports suggesting a bias towards maternal inheritance of the mutated chromosome 22 (Demczuk *et al.* 1995b, Ryan *et al.* 1997). VCFS has been shown to have an autosomal dominant mode of inheritance (Shprintzen *et al.* 1981, Williams *et al.* 1985). The phenotype of CTAF (Matsuoka *et al.* 1998) or Takao syndrome is very similar to VCFS (hyper nasal speech and mild mental retardation), which invites the suggestion that the two syndromes are actually differing manifestations — Oriental and Western — of the same phenotype (Neill 1987, Burn 1989). Opitz syndrome is the most recent addition to the CATCH22 family and was added by virtue of its craniofacial deformities. All Opitz anatomical abnormalities affect midline structures. In addition to affected cranial midline structures, patients present with hypospadias (Kimmelman 1982, Bershof *et al.* 1992). Heart defects have been described in only one Opitz patient (Jacobson *et al.* 1998).

What links all these syndromes together? I already mentioned that one of the central connections is a deletion (or translocation) in chromosome 22q11. The second link is that at least part of the phenotypes can be traced back to something gone awry in the development of the (cranial) neural crest. I can best show this by describing the defects behind each letter of the CATCH22 acronym.

III 3 2 1 Cardiac Defects

Abroad spectrum of cardiovascular anomalies is found in 22DS. Cardiovascular defects found in ADGS patients include interrupted aortic arch type B (IAA), persistent truncus arteriosus (PTA), pulmonary atresia (PA), ventricular septal defect (VSD), right aortic arch (RAA), and tetralogy of Fallot (TOF) (Payne *et al.* 1995) (figure 14). Associated with 22q11 deletions in general (*i.e.* including nonsyndromic patients) are also: absent ductus arteriosus, double outlet right ventricle (DORV), double inlet left ventricle (DILV), coarctation of the aorta, pulmonary valve stenosis (PVS), atrial

figure 14 Cardiac inflow, outflow and aortic arch arteries anomalies associated with neural crest maldevelopment. **A:** Normal heart. RA: Right Atrium; LA: Left Atrium; RV: Right Ventricle; LV: Left Ventricle. **B:** Double inlet left ventricle: both atria communicate with the left ventricle. **C:** Tricuspid atresia: tricuspid valve opening occluded. **D:** Straddling tricuspid valve. **E:** Persistent truncus arteriosus: the two main arteries do not separate because of truncal septation failure. **F:** Transposition of the great arteries: the aorta is connected to the right ventricle and the pulmonary artery to the left instead of vice versa. **G:** Double outlet right ventricle: both great arteries are connected to the right ventricle. **H:** Persistent truncus arteriosus over right ventricle: combination of DORV and PTA. **I:** Tetralogy of Fallot. **J:** Interruption of the aorta. **K:** Double aortic arch. **L:** Variable absence of the carotid arteries. **M:** Right aortic arch. **N:** Type B interrupted aortic arch with aberrant right subclavian artery. RS: Right Subclavian; LS: Left Subclavian; RPA: Right Pulmonary Artery; LPA: Left Pulmonary Artery; RCC: Right Common Carotid; LCC: Left Common Carotid.
Modified from Srivastava and Olson (1996) (A, E, F), Kirby and Waldo (1990) (B-D, J-H), Kirby and Waldo (1995) (G, H), Larsen (1993) (I), and Conley *et al.* (1979) (N).

septal defect, atrioventricular septal defect and transposition of the great arteries (Ryan *et al.* 1997) (figure 14). Monosomy of 22q11 is often seen in patients whom do not present with a DGS or VCFS phenotype but do have congenital heart disease (CHD) (Goldmuntz *et al.* 1993, Payne *et al.* 1995, Iserin *et al.* 1998). Velo Cardio Facial Syndrome has, as its name suggests, cardiac anomalies as part of its spectrum: VSD with or without RAA, TOF, coarctation of aorta, persistent ductus arteriosus, DORV, PS (figure 14). Cardiovascular malformations associated with CTAF are very similar: TOF, PA, RAA, sometimes PDA and often various anomalies in aortic arch artery derivatives (Momma *et al.* 1996).

From discussions in previous chapters we know that the cardiac neural crest from the third through sixth branchial arch contributes considerably to conotruncal and ventricular septation. Failure to migrate or develop properly in these areas results in PTA or VSD. At the time of septation there is still extensive looping of the heart and various malformations arise from problems in this process: e.g. DORV, DILV and PTA over right ventricle (= combination of PTA and DORV). Improper development of neural crest cells with respect to their contribution to aortic arches can result in the absence of aortic arches. Absence of aortic arch 3 or 4 results in IAA, while absence of arch 6 results in absent ductus arteriosus (the ductus arteriosus normally connects the pulmonary and aortic arteries).

In conclusion, CATCH22 patients are defective in parts of heart and vessel development, in particular those related to the third, fourth and sixth branchial arches.

III 3 2 2 Abnormal Facies

Both structurally and indirectly, the neural crest is of vital importance to the development of the face (bones, cartilage, muscle connective tissue), at least as studied in avian and amphibian embryos (le Lièvre and le Douarin 1975, Noden 1983a, b, 1984, 1988, Couly *et al.* 1993). Therefore we can be reasonably certain that the neural crest is involved in many of the facial features presented in DGS (short broad nose, a short philtrum, dysplastic low set ears, hypertelorism, up- or downwards slanting eyes, micrognathia), VCFS (broad nasal root with a prominent nose, retrognathia), CTAF (small mouth, short philtrum, hypertelorism, malformed auricles, low set ears, prominent ears, flat nasal bridge, narrow palpebral fissures and lateral displacement of inner canthi) and Opitz G (hypertelorism, prominent forehead, narrow palpebral fissures with epicanthal folds, dysphagia, stridor, laryngotracheal oesophageal clefts). Figure 15A shows some of the above mentioned facial features in cartoon form, while figure 15B is a photo of a VCFS patient.

A relevant feature in the context of head development, one not described by the CATCH22 acronym but often found in DGS patients, is deafness (Ryan *et al.* 1997). The crest not only contributes to the external ears, but also to the bones of the inner ear (see chapter II-3-3).

In conclusion, malformations of face, cranium and auditory system point to the neural crest as the common denominator.

III 3 2 3 Thymic Hypoplasia or Aplasia

In chapter II-3-4 I described the development of the thymus as the result of the interaction of endoderm and crest cells and how two thymus progenitors arise from the ventral aspect of the third pharyngeal pouch (receiving both structural and inductive contributions from the neural crest). Some CTAF and DGS patients are athymic, but most patients do have a thymus. However, it is often underdeveloped in DGS patients (Huber *et al.* 1967, Kretschmer *et al.* 1968). Experimental evidence shows that undersized thymi can be created by extirpating cranial neural crest in chicken (Bockman and Kirby 1984). DGS type neural crest related lymphoid deficiency distinguishes itself from other lymphoid deficiencies where the stem cells are missing or where they failed to interact with the endothelium (Bockman and Kirby 1984, 1985). Most patients have some form of immunocompromise, but the severity does not seem to correlate with any phenotypic feature (Sullivan *et al.* 1998). Immune deficiencies are seldomly the cause of premature death in DGS patients: the majority of deaths is related to congenital heart disease.

III 3 2 4 Cleft Palate

The V part of the VCFS phenotype is caused by problems in the development of the palate. VCFS patients suffer from velopharyngeal insufficiency, that is the inability to close the velopharyngeal portal (an opening between the soft palate and the pharynx), which affects normal speech (the patients have hyper-nasal speech). Cleft palate is also seen in DGS and sometimes in CTAF; Opitz patients often combine cleft palate with cleft lip. Cleft palate is a quite common birth defect and can be secondary to malformations of other parts of the face. Mouse knock-outs for genes that are not even expressed in the palate, such as *Msx-1* and *Hoxa2*, can result in cleft palates (Gendron-Maguire *et al.* 1993, Rijli *et al.* 1993, Satokata *et al.* 1994).

and this connecting them to the appropriate receptors is utilized by neural crest cells, as is separation of the left and right hemispheres. When these processes go awry, the following are seen: hypertelorism (excessive transverse or PTA) and the two variations: communication hemisphericity.

The palate forms from mesencephalic neural crest, so direct involvement of this neural crest population in cleft palate phenotype is very plausible. But, as palate development is dependent on the proper development of other parts of the face, such as the maxillae, the neural crest responsible for these parts of the face could be equally responsible.

III 3 2 5 Hypocalcemia

Just like the thymus, parathyroids depend on cranial crest cells for their formation. Since the parathyroids (and the thyroids) produce hormones that regulate systemic calcium levels, hypocalcemia can result from absence or improper activity of neural crest cells during parathyroid induction and development.

III 3 2 6 22q11 Aberrations

Deletions of the proximal long arm of chromosome 22q11 are among the most common structural chromosome abnormalities in man, with an estimated incidence of 1 in 4000 to 5000 newborns (Burn *et al.* 1993). Deletion and translocation breakpoints for some patients have been mapped with varying degrees of precision and some have been cloned. In the results section I will show the mapping within a several kilobase window of the breakpoints in the widely used CATCH22 cell lines GM05878 and GM00980, which have translocations in one of their chromosomes 22. GM05878 has a balanced $t(10;22)(q26.3;q11.2)$ (Cannizzaro and Emanuel 1985) and GM00980 cells miss one copy of 22pter-q11 owing to a translocation of 22q11-qter to 11q25 (replacing 11q25-qter) (Fu *et al.* 1976). The cell lines are from an (unaffected) father of a DGS patient and from a VCFS patient, respectively. The DGS child of GM05878 had an unbalanced variant of the translocation, deleting pter-q11. An interesting cell line that has received much attention is the ADU line from a DGS patient with a balanced $t(2;22)(q14;q11.1)$ (figure 21), which has been mapped on the molecular level (Budarf *et al.* 1995). Deletion through unbalanced translocation is not the rule in patients and balanced translocations are even more rare. Most deletions are

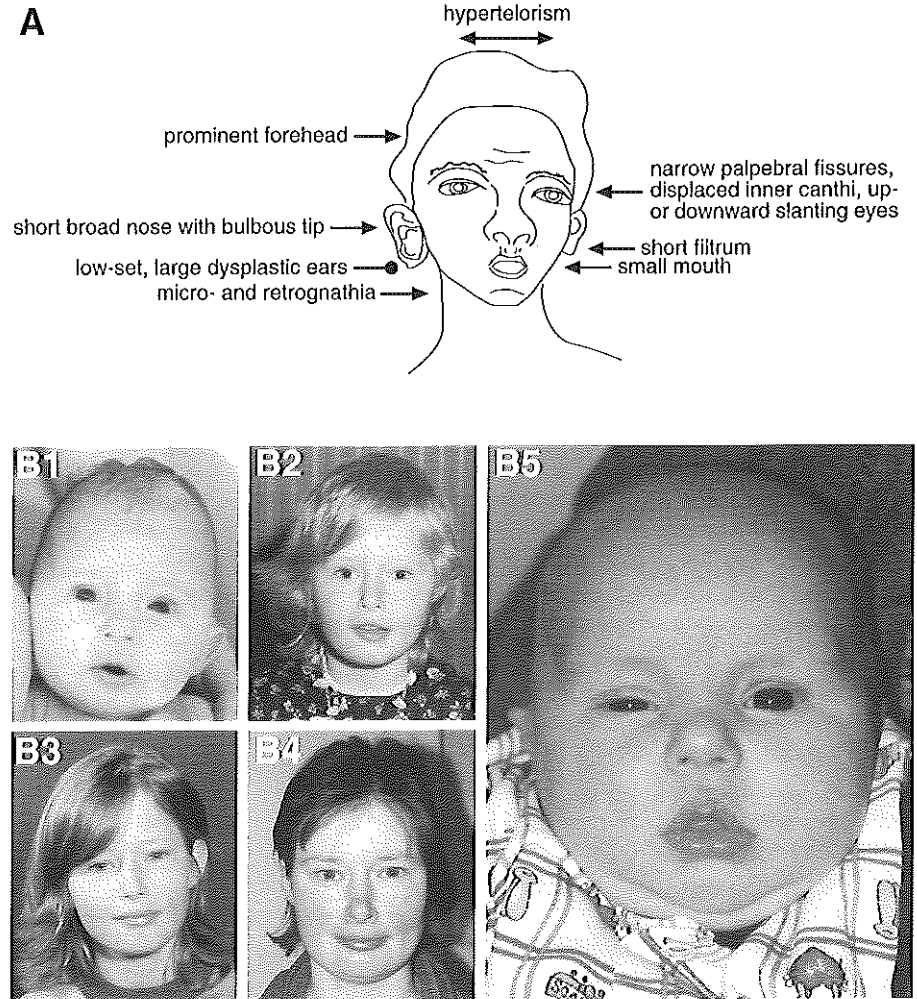


figure 15 Craniofacial features found in CATCH22 patients. **A:** Schematic illustration of facial deformities frequently observed in CATCH22 patients. **B:** Photographs of an actual VCFS patient. Propositus is shown at various life stages (baby till adult) (B1-B4). B5 is a photograph of her child, which also suffers from VCFS. Visible are the prominent, bulbous tip of the nose, narrow palpebral fissures with a slight upward slant, hypoplastic alae nasi and micrognathia. Overfolded helices are not visible on these photographs.
Published with permission. Photographs were kindly provided by Dr H. Westels.

interstitial, though still substantial in size at an estimated 2-3 Mb of DNA, usually including the chromosomal locus sc11.1B (Desmazes *et al.* 1993a, b, Halford *et al.* 1993c) (figure 21). Only a limited number of CATCH22 patients have an interstitial deletion smaller than the commonly deleted region (Levy *et al.* 1995, Dallapiccola *et al.* 1996, Kurahashi *et al.* 1996). Some DGS deletions are cytogenetically detectable (Wilson *et al.* 1992), but in those cases where this is not possible, molecular probes (figure 21) detect sub-microscopic deletions in more than 95% of the cases (Scambler *et al.* 1991, Carey *et al.* 1992, Driscoll *et al.* 1992a). A series of well characterized probes is available for this purpose (Carey *et al.* 1990, Fibison *et al.* 1990, Driscoll *et al.* 1992a, b, Aubry *et al.* 1993, Lindsay *et al.* 1993, Kurahashi *et al.* 1994, Mulder *et al.* 1995, Morrow *et al.* 1995). Using probes like these, some non-syndromic patients with conotruncal heart defects have been shown to have 22q11 deletions (Wilson *et al.* 1992b, Goldmuntz *et al.* 1993). 22q11 Deletions associated with cardiovascular anomalies can be an indication of associated (minor) extracardiac anomalies (Amati *et al.* 1995, Rauch *et al.* 1998). Probes are also being used to define (deletion)breakpoints and thus the shortest region of deletion overlap. The current region (SRD01) is defined between the centromeric breakpoint in DGS patient G (Levy *et al.* 1995) and the unbalanced translocation breakpoint in cell line GM00980. Recently, Kurahashi *et al.* provided evidence for the existence of a second critical region located in the distal part of the commonly deleted region (Kurahashi *et al.* 1996, 1997). The two SRD0s are mutually exclusive.

In conclusion, the vast majority of 22DS patients are hemizygous for part of chromosome 22q11. From patient to patient there is a wide variety in the size of the deletion.

III 3 3 A CLEAR LINK BETWEEN CHROMOSOME 22Q11 AND NEURAL CREST DEVELOPMENT

Experimentally, several of the 22DS phenotypes (disturbance of thymus, cardiovascular mal-development) can be recreated in animal models by deleting a specific part of the cranial neural crest, i.e. the cardiac neural crest (Bockman and Kirby 1984). Add to that the links between other CATCH22 phenotypes and neural crest, and I think it is not unreasonable to assume that the common etiology for the greater part of 22DS malformations is a disturbance in the development of the mesenchymal derivatives of the cranial neural crest. Various knock-out mice display phenotypes that are very similar to CATCH22 phenotypes. *Endothelin-1* (Kurihara *et al.* 1995a, b) and *Hoxa3* (Manley and Capecchi 1995) are the best known examples. More recently genes like *Rae28*, *ECE1*, *EDNRA*, *Patch*, *Pax3*, *Pax7* and *Pax9* have come into the picture. None of these genes, incidentally, are located on chromosome 22.

The variety of phenotypes seen in DGS patients could easily be explained with Thomas *et al.*'s description of theoretical variants of DGS: depending on which contiguous series of pharyngeal pouches and arches is affected, thirty-eight variants are possible (Thomas *et al.* 1987). Classical DiGeorge is defined as those cases where fourth arch and adjoining third and fourth pouch deriva-

tives are affected. The range of possible malformations can be extended even further were one to drop the contiguity condition. It is doubtful, however, that such a schema is biologically believable.

Having made the assumption that disturbance of neural crest development is causative for CATCH22 phenotypes, we can close the circle with the link to chromosome 22 deletions. We have seen in the previous chapter that most CATCH22 patients are deleted for chromosome 22q11, having either a large cytogenetic deletion or a smaller sub-microscopic deletion. Thus, considering the neural crest angle, we assume that 22q11 or its surrounds contains one or more genes essential for proper cranial neural crest development. That is the central theme of this thesis.

III 3 4 SYNDROME, DEVELOPMENTAL FIELD DEFECT, ANOMALY, MALFORMATION

In literature one can come across the term developmental field defect in connection with phenotypes like DiGeorge syndrome. A syndrome is defined as a set of concurring phenotypes caused by a known cause (Khoury *et al.* 1994), whereas a developmental field defect is defined as a defect that has a defined phenotype but a variable etiology (Lammer and Opitz 1986, Opitz and Lewin 1987). It can indeed be argued that in the case of DGS there is causal heterogeneity, though one can question how well defined the phenotype is. Cytogenetic abnormalities, maternal diabetes, and (ab)use of maternal alcohol (Amman *et al.* 1982), retinoid (Lammer *et al.* 1985) or bisdiamine can all cause DiGeorge syndrome or syndromes with very similar phenotypes (Belohradsky 1985, Lammer and Opitz 1986). Perhaps it depends on one's definition of cause. Is the actual teratogen the cause, or the crest mal-development induced by the various teratogens or genetic anomalies? At present the word syndrome is by far the preferred reference. It is mostly in the older literature where DGS is referred to as a developmental field defect because at the time it was not yet known that the majority of DGS patients have a 22q11 deletion or disruption. An international working group made definition and nomenclature recommendations for the various terms used to

IV CONCLUDING THE INTRODUCTION

Neural crest cells are very versatile: starting as neuroectodermal types, they give rise to neuronal, ectodermal, endodermal and mesodermal derivatives. In the regulatory sense there is more than just a hint of neuroectoderm background, though. As I described in the review of the regulation of crest cell migration, many molecules involved in neuronal guidance and proliferation act on neural crest cells as well. And not exclusively on the neurogenic sub population of crest cells. In many respects neural crest development parallels neurogenesis (Anderson 1989).

Since so many derivatives are affected in CATCH22 syndromes, we expect the gene or genes causing the phenotype to be quite high up in the hierarchy. Unless it turns out that we are dealing with a string of genes, each acting on a different sub-population of crest cells.

Quite some genes have been implicated in neural crest regulation, and the involvement of many has been confirmed. Studying syndromes with crest related phenotypes and associated chromosomal or genetic aberrations will certainly uncover many more.

describe the results of errors in prenatal development (Spranger *et al.* 1982). Their definition of a syndrome is "a pattern of multiple anomalies thought to be pathogenetically related and not known to represent a single sequence or a polytopic field defect". The definitions of sequence and polytopic field defect are "a pattern of multiple anomalies derived from a single known or presumed prior anomaly or mechanical factor" and "a pattern of anomalies derived from the disturbance of a single developmental field" respectively. These definitions describe patterns of morphological defects. With respect to CATCH22 the most relevant type of defect is malformation (malformity), described in Spranger *et al.* (1982) as "a morphologic defect of an organ, part of an organ, or larger region of the body resulting from an intrinsically abnormal developmental process".

Though in the context of DiGeorge the term syndrome is most widely used and probably the most accurate descriptor, in literature the more general term anomaly is sometimes used (Jablonski 1991, Khoury *et al.* 1994). Hence DGS is sometimes called DGA.

V
VI
VII
VIII

Results & Discussion

V INTRODUCTION TO THE EXPERIMENTAL WORK

Association of chromosomal deletions with a series of syndromes with overlapping phenotypes, suggest the presence of one or more genes in the deleted chromosomal area that are important for the (normal) formation of the affected tissues. The fact that most of the affected structures are (partially) derived from cranial neural crest or are dependent on interaction with it, makes it only a small step to suggest that some gene(s) in 22q11 play(s) a vital role in neural crest development. This is the central assumption for the work described in this thesis. Here I describe our work, trying to answer the following questions:

- what is the accurate position of known genes and markers in the region?
- can we find or predict new genes in the region and can we comment on their likelihood of being involved?
- what is the structure and expression pattern of one of the first candidate genes in the region ?
- is its expression pattern compatible with its candidate status?

The balanced $t(2;22)(q14.1;q11.1)$ translocation breakpoint in DGS patient ADU (Augusseau *et al.* 1986) was the major target of positional cloning in various laboratories. Critical region mapping and gene isolation strategies have been similar between most research groups. In general it entails testing markers (STSs) for linkage with 22DS (*i.e.* for deletion in patients), determining the linear order of different markers and delineating deletion boundaries. Confirmed markers are then used for the construction of contigs of genomic clones (cosmids, YACs, etc.). Cosmids are the most popular types of clones, and the process of iteratively isolating sets of overlapping cosmids and identifying overlapping cosmids by hybridizing back to the cosmid library to build and extend the contig in both directions is called cosmid walking. Following this step various techniques are in use to identify and isolate genes from the contigs in the critical region. cDNA library screening and exon trapping are the most

widely used techniques for the identification of potentially transcribed regions. Application of these types of experiments to the ADU breakpoint region led to the isolation of *DGCR3* and *DGCR5*. *DGCR3* is an open reading frame (ORF) of 260 amino acids (aa) that is interrupted by the translocation breakpoint (Budarf *et al.* 1995). However, no cDNAs were isolated and expression studies were only informative under low stringency conditions. This prompted Sutherland *et al.* (1996) to screen for adjacent coding sequences in the region. They identified a series of alternatively spliced transcripts from the *DGCR5* gene. The *DGCR3* sequence is embedded in an intron of *DGCR5* and has the same transcriptional orientation (but, see chapter VII-I-I). The *DGCR5* transcripts do not have an obvious protein encoding potential, but they could be functional RNAs. Current ideas about the SRDO places the ADU breakpoint circa 100 kb centromeric to the proximal deletion boundary of SRDO1 (Levy *et al.* 1995). Other genes then, located within this SRDO1, might be worth considering. Not only because of the unclear nature of the two genes (real gene?, pseudogene?, functional RNA coding gene?), but also because we do not know if only one gene is involved or more than one. For that reason it is important to get a clear view of all the genes in the region. Research groups the world over have actively searched for other potential candidate genes in the CATCH22 region, resulting in the transcription maps mentioned below.

The cosmid walking technique employed by several groups in Europe, the US and Japan was partly hampered by the apparent abundance of low copy repeats in the 22q region (Halford *et al.* 1993a). With large scale sequencing gaining momentum, methods maturing and resources becoming available, chromosome 22q11 became a focal point for one of the members of the HUGO (HUMAN Genome Organization) consortium. Dr Bruce Roe at the University of Oklahoma, in collaboration with the groups of Drs Beverly Emanuel and Marcia Budarf at the University of Philadelphia, produced the genomic sequence of a large, virtually contiguous, part of 22q11.2. Despite the difficulties caused by the repeats, the saturation cloning was at an advanced stage (Collins *et al.* 1995, Gong *et al.* 1996) and arguably all the genes that could be identified using conventional techniques had been located. A complete genomic sequence of the DGCR will aid in identifying more potential genes. We still know only little about the relationship between the CATCH22 syndrome and the genes in the DGCR. It is very well possible that any gene in the commonly deleted region that has some developmental effect can contribute to the phenotype when hemizygous. Therefore it is important to identify all genes in the region and gain information about their role in embryonic development. During my project two transcription maps of SRDO1 were reported (Gong *et al.* 1996, Lindsay *et al.* 1996). Eighteen genes were mapped to the commonly deleted region of chromosome 22q11 (Winqvist *et al.* 1991, Aubry *et al.* 1992, Aubry *et al.* 1993, Halford *et al.* 1993b, c, Kelly *et al.* 1994, Yagi *et al.* 1994, Demczuk *et al.* 1995a, Heisterkamp *et al.* 1995, Kurahashi *et al.* 1995, Lamour *et al.* 1995, Wadey *et al.* 1995, Demczuk *et al.* 1996, Gong *et al.* 1996, Kedra *et al.* 1996, Lindsay *et al.* 1996, Pizzuti *et al.*

1996, Rizzu *et al.* 1996, Sirotkin *et al.* 1996, Sutherland *et al.* 1996, Pizzuti *et al.* 1997). Sequences homologous to the 3' UTR of the human *DISHEVELLED 1* gene were mapped just centromeric to the *HIRA* gene (Pizzuti *et al.* 1996). It is unclear whether these sequences belong to a gene or are part of a pseudogene. Little is known about the expression of the CATCH22 candidate genes in adult and embryonic tissues. A 22DS candidate gene should at the very least be expressed embryonically, so it will be important to establish accurate expression patterns of candidate genes, and study in detail embryonic expression where applicable. A candidate gene is expected to be expressed in tissues relevant to the syndrome, *i.e.* cranial neural crest, or cells interacting with the crest. Even if it is expressed during development, a gene with expression only found in the tip of the tailbud would not seriously be considered a candidate gene.

Before we reached the stage of studying a candidate gene, we produced a low resolution map of the region, mapping locus D22S183 in relation to various translocation breakpoints and other markers in the DGCR (Mulder *et al.* 1995). One of the tools employed for this purpose was the construction of hybrid cell lines (hamster-human hybrid cells containing wild type or deleted human chromosome 22). Hybrid cell lines allowed us to map probes and study the extent of deletions in individual patients. We enhanced the low resolution mapping by using its results to construct several cosmid and YAC contigs in the 22q11 region with the cosmid walking technique mentioned earlier. Since the cosmids we used were sourced from the same Lawrence Livermore Chromosome 22 library as the majority of the clones sequenced by Dr Roe, I was able to fully integrate our contigs into the sequenced contigs. In collaboration with the Emanuel and Budarf groups I analyzed pre-released genomic sequence of the DGCR. Using a combination of biocomputing and experimental approaches, I set out to identify transcription units in the sequence dataset that are evolutionary conserved and developmentally regulated. From the sequences in the database, I constructed five contigs from 22q11 sequences and searched for expressed and conserved sequences by comparison with EST, yeast and fly sequence databases. In the Results and Discussion section I report the position and genomic structure of twenty-seven genes within the five contigs; of these twenty-seven genes twelve were known and fifteen were unknown at the time of analysis. Results were communicated to Philadelphia, where experimental analysis was started. Amongst the new genes I find *RANBP1*, a *SEPI1/DIFF6* homolog, the gene for a novel Glutathione/Thioredoxin Reductase like protein, a probable Catenin family protein gene and the human homologs of mouse *T10* and *TBX1*. A number of new genes display no homology to published sequences so the possible function of their products can not even be speculated on. X-grail analysis helped confirm new genes and suggested possible extra exons of incomplete genes. Detailed analysis of the alignment of cDNA sequences (ESTs) with the genomic DNA sequence reveals potential alternatively spliced mRNAs for several genes. For additional and updated analysis and advanced displaying of the annotated genomic sequence I used the ACeDB software suite.

As an initial step to determine the expression pattern of genes from the CATCH22 SRDO1 we chose to analyze the expression pattern of *HIRA*. The *HIRA* gene (*DGCR1*) was originally isolated as *TUPLE1* by Halford *et al.* (1993b). Lamour *et al.* isolated human cDNAs containing the complete *TUPLE1* sequence with 621 additional nucleotides in the ORF (Lamour *et al.* 1995). They renamed the gene *HIRA* because the most significant peptide homology found at that time was with Hir1p and Hir2p, two histone gene repressor proteins from the yeast *Saccharomyces cerevisiae* (Sherwood *et al.* 1993). The predicted HIRA protein is characterized by the presence of seven N-terminal WD40 repeats, two bipartite nuclear localization signals and a penta-Q stretch (frequently seen in transcription factors). The gene product has been implicated in transcriptional regulation based on its homology to gene regulators like yeast Hir1p, Hir2p and Tup1p (Sherwood *et al.* 1993, Promisel Cooper *et al.* 1994). I describe the identification of cDNA clones of the murine homolog of *HIRA* and studies on the expression of *Hira* during mouse embryogenesis. The sequence of the predicted protein is homologous with the p60 subunit of human Chromatin Assembly Factor 1 (CAF1A) (Kaufman *et al.* 1995). On the basis of sequence homologies with Tup1p, CAF1A, Hir1p and Hir2p, we propose that HIRA is involved in the assembly of chromatin, either by interacting with histones or by regulating their genes. Our results demonstrate that *Hira* mRNA is ubiquitously present from early developmental stages through adulthood. Raised levels of mRNA are detected in the neural folds, pharyngeal arches, circumpharyngeal neural crest and limb buds (Wilming *et al.* 1997). We think that HIRA is normally involved in several aspects of neurogenesis, pharyngeal arch development and limb development by regulating genes that control these processes.

VI LOW RESOLUTION MAP: CONSTRUCTION OF FOUR COSMID CONTIGS ON 22Q11

VI 1 ANONYMOUS SINGLE COPY PROBE NB84 LOCATES TO 22Q11

Single copy probes were assigned to the region 22pter-q11 (van Biezen *et al.* 1993). We were interested in gaining a starting point for a chromosomal walk across the DGCR, and to that end we tested these anonymous single copy probes for their localization within the DGCR. Source material for the DGCR deleted and normal chromosomes 22 consisted of human-hamster hybrid cells. Fibroblasts from two unrelated DGS patients were fused with thymidine kinase deficient Chinese hamster cells and hybrids retaining one chromosome 22 were selected.

By Southern analysis with probe HP500 or FISH analysis with HP500 cosmid sc4.1, we determined the presence of a deleted or wild type chromosome 22. HP500/sc4.1 is a probe for locus D22S134, which is hemizygous in 95% of normal DGS patients (Carey *et al.* 1992). We were able to assign probe NB84 (locus D22S183) (Lekanne Deprez *et al.* 1991) to the deletion of both patients (figures 16-18). This warranted further research into the nature of NB84 and

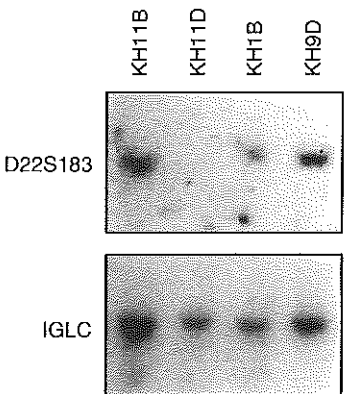


figure 16 Autoradiogram of a Southern blot of EcoRI-digested DNA from human (DGS)-Chinese hamster hybrid cell lines hybridized sequentially with probe NB84 (D22S183) and a probe from the immunoglobulin lambda light chain complex (IGLC) as a chromosome-22-specific control. Cytogenetic analysis had shown one chromosome 22 in each hybrid cell line. Hybrid KH11D carries the deleted chromosome 22; the other hybrids have retained the normal chromosome 22. From Mulder *et al.* (1995).

Figure 16: Southern blot analysis of EcoRI-digested DNA from human (DGS)-Chinese hamster hybrid cell lines. The probe NB84 (D22S183) and the immunoglobulin lambda light chain complex (IGLC) were used. The figure shows a 2x4 grid of bands. The top row is labeled 'D22S183' and the bottom row is labeled 'IGLC'. The columns are labeled 'KH11B', 'KH11D', 'KH1B', and 'KH9D'. Each column shows a pair of bands in the IGLC row and a single band in the D22S183 row.

figure 17 Quantitative Southern analysis of PstI-digested DNA from two patients (lanes 3, 4) and a control subject (lanes 1, 2; with a double amount of DNA in lane 2). The blot was hybridized simultaneously with probe NB84 (locus D22S183) and a probe for the B-globin locus as a control. Comparison of the signal intensity produced by the test probe and the control probe shows that one patient (lane 3), who was diagnosed as having VCFS, is hemizygous for D22S183. The other patient (lane 4) is heterozygous for the PstI polymorphism in D22S183 and therefore has no deletion at this locus. This patient has congenital heart disease and minor facial dysmorphisms. From Mulder *et al.* (1995).

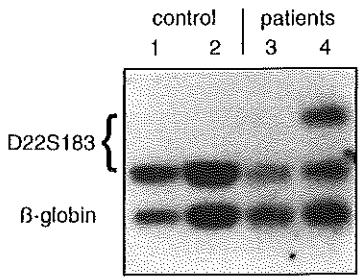
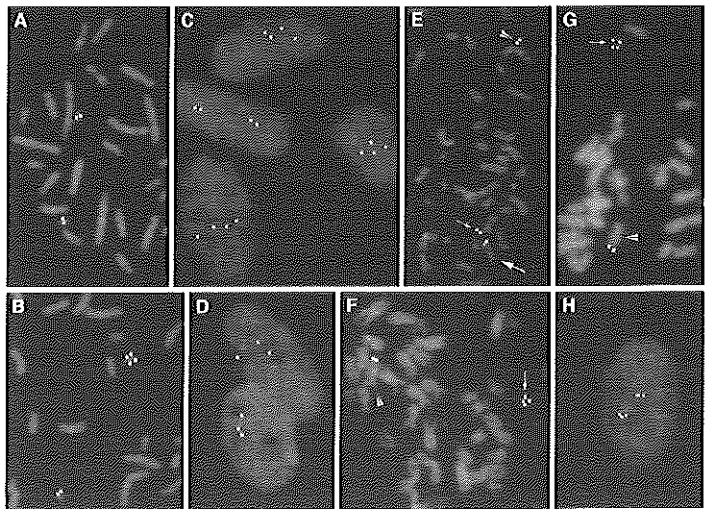


figure 18 (this is a grayscale thumbnail of the original for quick reference; see full size color picture on the foldout poster on the back cover)

FISH analysis of metaphase chromosomes and G1 interphase nuclei. Probes are colored either in green (fluorescein) or in red (rhodamine/Texas Red) and the DNA is counterstained in blue with DAPI. Fluorescence was registered using a triple band filter set. **A:** Partial metaphase spread from lymphocyte culture of normal control subject, hybridized with cosmid M51 (D22S183) labelled in red. Both chromosomes 22 show a signal in the q11 region. **B:** Partial metaphase spread from a DGS patient, hybridized with M51 (green) and the control probe M78 (red). M78 produces a signal on both chromosomes 22 (at q13), but M51 on only one indicating hemizyosity. **C, D:** G1 interphase nuclei from fibroblast cultures of a normal control subject (C) and a DGS patient (D), hybridized with M51 (green) and the chromosome-22-specific control probe M78 (red). Both M78 and M51 produce two signals per nucleus in the control (C), whereas in the DGS nuclei (D) M78 produces two signals but M51 only one. **E:** Partial metaphase spread from the balanced translocation cell line GM05878, hybridized with M51 (red) and sc4.1 (green). In the normal chromosome 22 (arrowhead), the signals of M51 and sc4.1 overlap, resulting in a white signal. The derivative 22 (small arrow) shows a red signal for M51 and the derivative 10 a green signal for sc4.1, indicating that the translocation breakpoint is between the two loci. **F:** Partial metaphase spread from the unbalanced translocation cell line GM00980, hybridized with M51 (green) and M78 (red). The normal chromosome 22 (arrow) shows proximal and distal signals for M51 and M78, respectively. The derivative 11 (arrowhead) shows a signal for M78 but not for M51, indicating that M51 is deleted and maps centromeric to the translocation breakpoint in this cell line. **G:** Partial metaphase spread from cell line GM00980, hybridized with M25 and the control probe pH17, both in red (the two probes were supplied by the manufacturer, both labelled with digoxigenin, making two-color hybridization impossible). The normal chromosome 22 (arrow) shows proximal and distal signals for M25 and pH17, respectively. On the derivative 11 (arrowhead), only the distal signal is present, indicating that M25 is deleted and maps centromeric to the translocation breakpoint. **H:** G1 interphase nucleus from the fibroblast culture of a normal control subject. Triple hybridization with cosmids ch748 (red), M51 (green) and sc4.1 (red). Both triplets show the order red-green-red. In one triplet, the green signal partially overlaps one of the red signals. From Mulder *et al.* (1995).



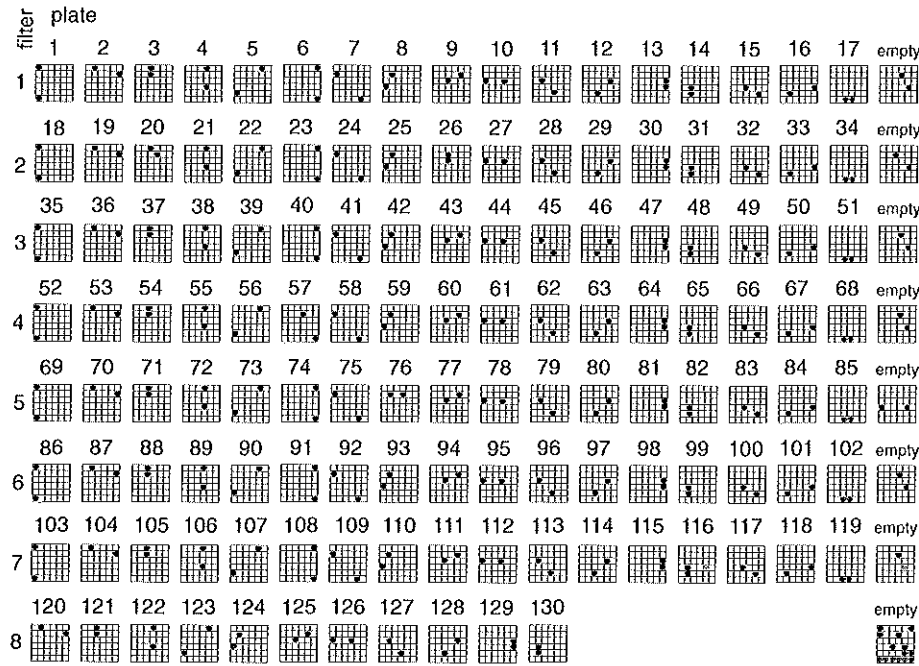


figure 19 Coding diagram used for the blotting of high density cosmid filters. Each row represents a filter containing clones from up to seventeen 96 well library plates. Plates were blotted (with a 96 pin robot) in duplicate on each filter at positions indicated here with black dots. A pair of dots was left empty in each cell for easy positional reference. On an autoradiogram the filter number and cell coordinate are easily determined assuming proper labelling practice. Together with the unique combination of the configuration of a signal pair and its position within the cell identifies the originating plate (1-130). This unambiguously identifies the well containing the clone of interest.

its position relative to previously published markers and loci in and surrounding the DGCR.

Radiolabelled NB84 was hybridized to a Zoo blot containing *EcoRI* digested genomic DNA from various animal species and yeast. We only saw a clear hybridization signal with human and Rhesus monkey DNA, indicating conservation of D22S183 sequences in primates. Weak signals are present in DNA from other mammalian species as well (not shown).

Conservation, though limited in this case, can indicate the presence of coding sequences. This was tested by hybridization of NB84 to a Northern blots containing polyA⁺ RNA from various adult human organs and a fetus. We did not observe any signals on these Northern blots, so the likelihood of the presence of expressed sequences at D22S183 is low. We sequenced the 1 kb insert of NB84 and did homology searches through sequence databases. No significant homologies with known expressed sequences were found.

VI 2 GRIDDED CHROMOSOME 22 COSMID LIBRARY

Since NB84 is located in the DGCR but does not appear to represent coding sequences, we used the probe to identify genomic clones representing the deletion. For this purpose we decided to use the gridded human chromosome 22 cosmid library available from the Lawrence Livermore Laboratories. This library is supplied as frozen bacterial stocks in 130 numbered 96-well dishes. The advantage of gridded libraries is their ease of replication and their standardization. Every well (= cosmid clone) has a unique coordinate, greatly facilitating communication between different researchers using the same library. The LLL chromosome 22 library is a well established one, used by many groups in the CATCH22 and chromosome 22 fields. If we would decide to sequence the cosmids of interesting contigs, the results could be tied in immediately with a large body of existing data, something not possible if disparate libraries are used.

A Beckman BioMek robot at the State University at Leiden (Netherlands) was used to plate out the colonies in a gridded array on nylon membranes. Through the custom software written by W. van Loon we were able to plate in a 6x6 array, i.e. 96 blocks of 36 colonies per 8x12 cm membrane. Such a high density of colonies causes two problems that had to be addressed. Firstly, it is virtually impossible to determine the exact coordinate of a positive signal on the membrane, especially considering positional inaccuracies caused by slight deviations from true of individual blotting "fingers" of the robot due to movement or bends. Secondly, due to the small size and high density of the DNA spots the distinction between true signals and random background spots will often be impossible to make. To circumvent the latter problem, I designed a system where each clone is represented twice in the 6x6 array. Thus only signal pairs within one block are considered true positives. To alleviate the first problem, we took advantage of the pairing system by transferring colonies to the membranes in such an order that each colony pair has a two-dimensional arrangement that cannot be mistaken for any of the other 17 pairs. Pencil marks between blocks indicate the position of the blocks on the membranes. The shape of the signal

pair and the position within the block unmistakably reveals the coordinate of the original clone. **Figure 19** shows the decoding table used to decipher the patterns. A typical result of a hybridization of the library filters is shown in **figure 20**.

Filters with gridded YACs are available from various sources. These offer the same advantages as the above filters and come ready to hybridize. We used gridded YAC library filters representing various libraries to find clones connecting isolated cosmid contigs.

VI 3 NB84 RECOGNIZES FOUR COSMIDS FROM TWO SITES ON CHROMOSOME 22

With NB84 we screened the gridded human chromosome 22 cosmid library L22NC03. This resulted in the isolation of four cosmids. Their coordinates are 19B9 (lab code M47), 33D9 (M51), 38A6 (M56) and 83C5 (M67) (see also **table 4**). The *EcoRI* restriction patterns were compared, showing that M51 overlaps with M47 and M67 (referred to as the M51 group). M56 does not overlap and has three *EcoRI* bands weakly hybridizing to NB84, as opposed to the one band with which NB84 strongly hybridizes in M47, M51 and M67. Clearly NB84 has some weak similarity to sequences on M56. FISH on interphase chromosomes confirmed that M51 and M56 represent separate loci on chromosome 22. As the signals overlap on metaphase chromosomes, the distance between the two loci is expected to be in the range of several hundred-thousand basepairs.

VI 4 M51 IS LOCATED CENTROMERIC OF M56 BUT BOTH ARE LOCATED WITHIN THE DGCR

To determine whether M51 and M56 are located in the DGCR and what their position relative to known reference points (breakpoints, probes) is, we performed FISH experiments with various probe combinations and cell lines.

We used cell lines from individuals with translocations in the DGCR: ADU, GM05878 and GM00980. The ADU cell line is from a patient with a balanced t(2;22)(q14.1;q11.1) translocation;

† fragment length polymorphism showed that the marker was missed on one of their chromosomes. Having a DNA probe from the DGCR available meant that we could use additional cosmid impressions for DGCE. We remember to take one or more copies of DGCE pictures, there is no need to use them to identify

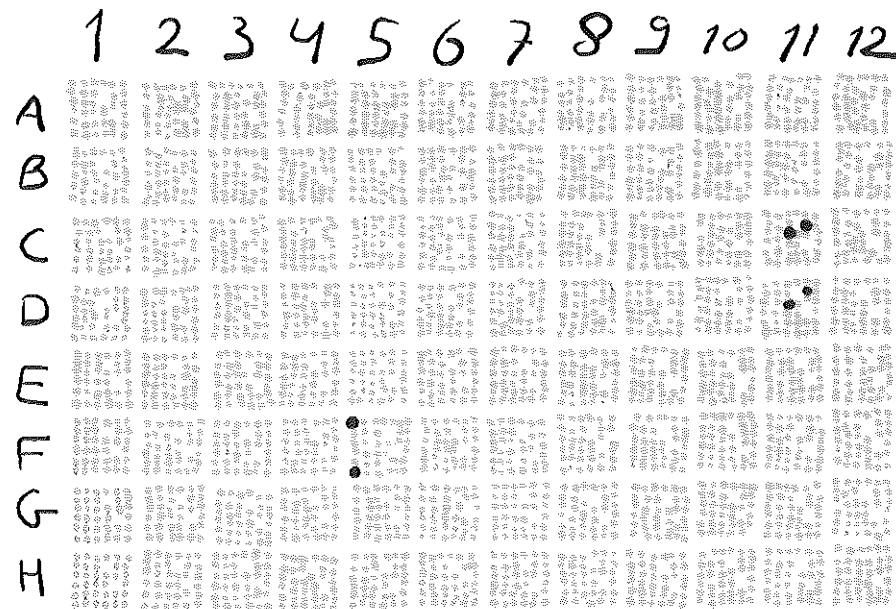


figure 20 A simulation of a typical autoradiogram resulting from exposure to filter 4 hybridized with ³⁵S-labelled vector DNA (yielding the weak homogenous background signal for easy reference) and ³²P-labelled H56 probe (the three black signal pairs). The detected overlapping clones are (see diagram figure 19) cosmids H202 (5215) H203 (60C11) and H204 (60D11).

the patient suffers from mild DiGeorge. GM05878 cells have a t(10;22)(q26.3;q11.2) genotype and are from an unaffected father of a child which suffers from DiGeorge. The child has an unbalanced variant of the same translocation as the father, causing 22pter-q11 monosomy (Kelly *et al.* 1982). Finally, GM00980 cells are from a VCFS patient with an unbalanced translocation t(11;22)(q25;q11), deleting 22pter-q11 (Fu *et al.* 1976). Two color FISH (**figure 18**) with various combinations of markers M51 (D22S183), M56, cH748 (*HIRA/TUPLE1*), N25 (D22S75) and *ZNF74* suggested the following order of markers (breakpoints in bold type): centromere - sc11.1A - **ADU** - *HIRA* - M51 - N25 - **980** - M56 - **5878** - sc4.1 - sc11.1B - *ZNF74* - telomere (**figure 21**). This shows that both M51 and M56 are located in the DGCR, M51 proximally from M56. Sequence analysis several years later shows that the position of N25 is incorrect as it is located between ADU and *HIRA*.

With RFLP and quantitative Southern hybridization we were able to confirm on of the outcomes of the FISH experiments, namely that M51 and M56 are deleted in DGS patients (**figures 17, 22**).

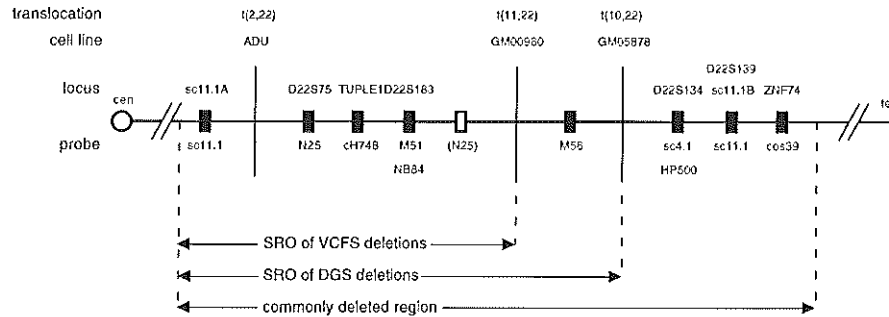


figure 21 Diagram of the DiGeorge chromosomal region at 22q11 showing the relative order of loci (black boxes) and translocation breakpoints (vertical lines) studied. The legend on the left indicates the type of label used on that line (probes can be cosmids or plasmids). Drawing is not to scale. Outline box indicates the incorrect position of N25 as initially published. cen: centromere; tel: telomere. Modified from Mulder et al. (1995).

VI 5 M51 COSMID CONTIG

We probed the cosmid library with cosmid M51, initiating a cosmid walk in the DGCR to isolate a contig of overlapping cosmids representing part of the DGCR.

M51 hybridization yielded five new cosmids shown in **table 4**. Restriction and hybridization analysis showed that M101, M107 and M108 overlap extensively, with M101 close to the M51 group and M108 extending outwards. The three cosmids extend the contig in one direction only. There are no cosmids overlapping with the other end of the M51 group. Repeated attempts with M51 and M67 probes did not yield any overlapping cosmids other than the known ones. In an attempt to bridge the gap on that side of the contig, we hybridized M67 to a human CML0 total DNA phage library (constructed and kindly provided by Dr C. Troelstra). This experiment did not yield overlapping clones. M108 extending outwards most, we chose this cosmid as the probe for the next step.

This yielded seven cosmids (**table 4**). Detailed fingerprinting showed that for only five of these, M124, M127, M129, M130 and M131, the restriction digest pattern could be resolved. M140 and M142, the remaining two cosmids, overlap as a pair, but could not be placed in the contig without serious conflicts on the level of restriction fragments (*EcoRI*, *BamHI* and *HindIII*). Nevertheless, both on hybridization level and on restriction enzyme level there is a considerable overlap between the consensus contig and M140 and M142.

FISH with M108 shows two signals on metaphase chromosomes 22: one in the DGCR and one close to the centromere on the p-arm. An M127 probe only hybridizes to a region in the DGCR. In the KHI 1D cell line (a hybrid cell line containing a deleted chromosome 22, lacking D22S185 but

containing D22S184) this signal is not present, indicating that, as expected, M127 represents a locus inside the DGCR. The M108 sequence is wholly contained within the combination of M47 and M127, neither of which show double FISH signals, making it difficult to explain why M108 does. M131 also shows two FISH signals. One of the signals overlaps with the M108 DGCR signal, the other signal is also located on chromosome 22, but pericentric. FISH with M140 resulted in signals on all acrocentric chromosomes.

Since coding regions frequently have a CpG island 5' of the coding region, we tested M67, M108, M127, M131 and M140 for the presence of restriction sites rich in GC by digestion with *BssHIII*, *NotI* and *SacII*. M67 is not digested by any of the enzymes. M127 and M131 have three *BssHIII* and one *SacII* site, whereas M108 has two *BssHIII* and one *SacII* site. *NotI* sites are not present in these cosmids. M140, the cosmid that clearly overlaps with the M51 contig but can not be faithfully integrated into that contig, displays many sites for the three enzymes. We conservatively estimate that M140 has fifteen *BssHIII* sites, eight *NotI* sites and fifteen *SacII* sites. Clearly this cosmid contains a CG-rich region. This sequence stretch must be located outside the overlap with the M51 contig, since the contig cosmids do not have more than a few *BssHIII* and *SacII* sites. The high GC content may explain the FISH results.

VI 6 M56 COSMID CONTIG

As mentioned earlier, the NB84 hybridization to the cosmid library yielded the M51 group, but also one isolated cosmid: M56. Since M56 turned out to represent DGCR sequence as well, we used this cosmid as a second starting point for a cosmid

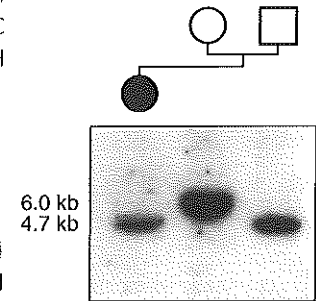


figure 22 Autoradiogram of a Southern blot of *PstI*-digested DNA from a proband and her parents, hybridized with probe NB84 (locus D22S183), which reveals a polymorphism in this family. The mother and father are homozygous for the 4.6 kb and 4.0 kb allele, respectively. The proband failed to inherit the maternal allele. from Mulder et al. (1995).

exposed to Southern hybridization with labelled sheared human placenta DNA and vector DNA. Non-hybridizing bands were considered single copy and isolated. Probes SCP1, SCP2 and SCP3 are 290 bp, 220 and 700 bp *Pst*I fragments from M108, respectively. SCP4, SCP5 and SCP6 are isolated from the M56 contig and are a 620 bp *Pst*I fragment from M204 and 650 bp and 1050 bp *Pst*I fragments from M202, respectively. We hybridized SCP1-6 to Southern blots containing restriction digests of cosmids from the M51 and M56 contigs (not shown).

VI 8 HIRA/TUPLE1 AND HD7K CONTIGS

Earlier we determined the position of *HIRA* (*TUPLE1*), a cDNA from the DGCR, relative to M51 and M56. In an attempt to reach the M51 contig from the *HIRA* side, we hybridized the *HIRA* cDNA to our cosmid library. Since *HIRA* is thought to be coded for by an estimated 150 kb genomic DNA (Halford *et al.* 1993b), an expected large number of cosmids (twenty-six) hybridized to the cDNA probe (table 4).

A third starting point for a DGCR cosmid contig was provided by probe HD7K. This probe was assigned to the DGCR by Wadey *et al.* (1993). Twenty-one cosmids were confirmed positives (table 4).

VI 9 YACS

To bridge the gaps and circumvent the 'road blocks' we encountered during the cosmid walk, we screened ICRF gridded human YAC library #900 with NB84. This yielded one positive clone, Y900B0267. This YAC is approximately 610 kb and does not contain *HIRA*.

A human *HIRA/TUPLE1* probe recognized a different YAC clone: Y900H05113. The 2.6 kb cDNA probe recognizes *Eco*RI bands of approximately 3, 5.5 and 12 kb on this YAC.

A 3' *HIRA/TUPLE1* probe (350 bp *Ssp*I-*Cla*I) we used to probe H05113 and *HIRA/TUPLE1* positive ICI human YACs 10BE5 (245 kb), 30FH10 (290 kb) and 37AF5 (200 kb) (a kind gift from Dr P.J. Scambler) revealed that H05113 and 37AF5 are devoid of this 3' sequence. NB84 hybridizes to these YACs, so all three ICI YACs contain the D22S183 locus.

For the purpose of transgenic mouse experiments we set out to isolate large mouse genomic fragments representing the equivalent of the DGCR. Mouse *HIRA/TUPLE1* probes on ICRF mouse YAC libraries #902, 903, 909 and 910 resulted in the isolation of one YAC. This 1 Mb clone, Y903E1213, has 5, 6 and 11 kb *Eco*RI fragments containing sequences homologous to the probe. FISH suggests that this YAC is chimeric because signals on chromosome 16 and 8 are observed.

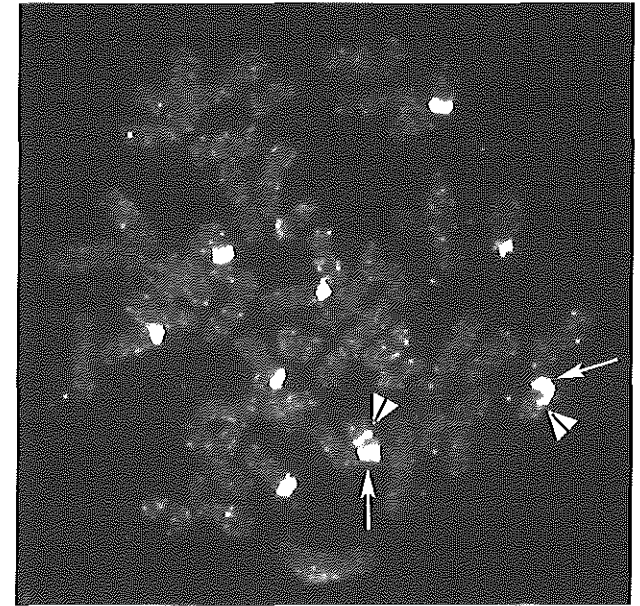


figure 23 FISH analysis of normal metaphase chromosomes with cosmid M80 (arrow) and a chromosome 22 specific control probe (arrow head) showing strong hybridization to acrocentric chromosomes besides chromosome 22.

VI 10 cDNA LIBRARY SCREENINGS WITH CONTIG COSMIDS: M51 AND M56 CONTIG COSMIDS DO NOT CONTAIN ABUNDANT OR COMMONLY EXPRESSED SEQUENCES

The reason for the construction of cosmid contigs was to have manageable chromosome 22q11 fragments available that can be tested for the presence of expressed sequences, either human or mouse. We hybridized M51 to a 26 week human fetal brain cDNA library and a human testes cDNA library. No positive clones could be detected.

M67 and M108 were tested by low stringency hybridization to a mouse fetal cDNA library and hybridization to the human fetal

cDNA library. M56 and M202 were only tested on the human library. We were not able to isolate positive clones from these screenings.

VI 11 DISCUSSION

Using cosmid probe M51, we have determined the map position of the anonymous marker locus D22S183 relative to a number of translocation breakpoints and other known loci. D22S183 maps within the SRO. A second locus, recognized by cosmid M56 (which contains sequences with weak homology to D22S183), is located more distally: just outside the SRO. A positional map is presented in **figure 16**. Several cosmid contigs, totaling ≈ 400 kb, were established. Progress in extending the two main contigs, M51 and M56, was halted by the failure of end cosmids to hybridize with any new cosmids, or to hybridize with unique overlapping clones. Chimerism or low-copy repeats may be responsible for this behavior. A later chapter (**VII-7**) will discuss this further.

From the initial course mapping, we estimated D22S183 to be less than 200 kb from *TUPLE1/HIRA*. Precise mapping (discussed below) shows that D22S183 is located 40 kb from the 5' end of *HIRA* and 140 kb from the 3' end.

In the absence of evidence that deletion or mutation of any of the known genes in 22q11 is sufficient and/or necessary to cause CATCH22, the identification of other genes in or around the SRO remains an important issue. In this respect, D22S75 and D22S183 may be relevant loci. I performed a detailed analysis of the region to assess this.

VII HIGH RESOLUTION MAP: MAPPING OF NEW GENES IN THE 22Q11 REGION

Dr B. Roe's laboratory in Oklahoma sequences genomic clones from the CATCH22 region for Dr B. Emanuel's group in Philadelphia. I cannot illustrate the advantage of gridded libraries better than with this sequencing project: we were able to integrate our cosmid contig data with the sequence data thanks to the fact that most sequenced clones were from the same LLL library we used.

As the sequences from these large scale sequencing projects were made available, I analyzed 22q11 DNA sequences with a bioinformatics approach. 1.65 Mb of 22q11 sequence was fetched from the EMBL database (AC000067-AC000071, AC000073-AC000080, AC000082-AC000095, AC000097, L77569, L77570). I constructed a FastAble 22q11 database from these sequences and determined the overlap between the clones. Contiguous sequences were assembled into five consecutive non-overlapping contigs: from centromere to telomere contig 1 through contig 5. The total length of the non-redundant sequence amounts to approximately 1.2 Mb. Individually the contigs are sized as follows: 178 kb (contig 1), 47 kb (contig 2), 421 kb (contig 3), 343 kb (contig 4) and 207 kb (contig 5).

Through FastA alignment I tested our 22q11 sequence database for the presence of DGS candidate genes and known 22q11 transcription units. *DGCR2/IIDD/LAN/DGSC* (Demczuk et al. 1995a, Wadey et al. 1995; X83545), *DGCR5* (Sutherland et al. 1996; X91348), *DGCR3* (Budarf et al. 1995; S579494), *DGSA*, *DGSB*, *DGSD*, *DGSE*, *DGSF*, *DGSG*, *DGSH*, (Gong et al. 1996; L77559, L77561-L77565, L77571) *DGSI/ES2* (Gong et al. 1996, 1997; L77566), *CTP/DGSJ* (Kaplan et al. 1993, Helsterkamp et al. 1995; U25147), *CLTCL1/CLTD/CLH-22/DGSK/ES3* (Kedra et al. 1996, Long et al. 1996, Sirotkin et al. 1996; U41763), *HIRA/DGCR1/TUPLE1* (Halford et al. 1993b, Lamour et al. 1995; X89887), *UFD1L* (Pizzuti et al. 1997; U64444), *COMT* (Brahe et al. 1986, Winqvist et al. 1991; M58525) and *GP1BB* (Kelly et al. 1994, Yagi et al. 1994; J03259) genes are present in our 22q11 database. Mouse genes mapping to the syntenic region of human chromosome 22q11 (mouse chromosome 16) were also compared to our 22q11 database, resulting in the mapping of the human homologs of mouse *T10* (Halford et al. 1993c; X74504) and *Tbx1* (Bollag et al. 1994; U57327). The linear order of the genes, including new genes discussed later, is shown in an overview (**figure 24**). New anonymous genes are labeled DiGeorge Deletion Transcript (DDT). I do not find genomic sequences representing the (potential) 22q11.2 genes *DVL22* (Pizzuti et al. 1996; U46462), *ZNF74* (Aubry et al. 1992; X71623), *UBE2L3* (Moynihan et al. 1996), *DGCR6* (Demczuk et al. 1996; X96484), *HCF2* (Blinder et al. 1988, Herzog et al. 1991; M12849, M19241), *GSTT1* (Webb et al. 1996; X79389), *GSTT2* (Tan et al. 1995; L38503),

MIF (Budarf *et al.* 1997; M25639, L19686) and *LZTR* (Kurahashi *et al.* 1995; D38496). Individual contigs were analyzed for the presence of expressed and conserved sequences by database comparison of 40 kb fragments of contigs with EST, yeast translated and *Drosophila* translated databases. Contig fragments from regions void of conserved or expressed sequences were compared to the translated EMBL database. With these approaches I discovered fifteen new potential transcription units in contigs 3, 4 and 5 (I do not find new transcription units in contigs 1 and 2). These genes, as well as the established genes, are discussed in the following paragraphs. Discussion follows the linear order on the 22q11 map from centromere to telomere (figure 24). We analyzed the contigs with X-grail to locate potential promoters, CpG islands and polyA signals. X-grail was also used to predict exons. The potential genomic range of new genes for which we have little or incomplete information (i.e. only few ESTs or only partial peptide homology) is predicted using the X-grail suggested exons.

I map chromosomal markers with known sequence to the 22q11 region and establish their precise linear order (D22S121, -183, -185 (Budarf *et al.* 1996), -553 (gdb 335921), -609 (gdb 335977), -931-933, -941-944 (gdb 434655, 434658, 434661, 434689, 434693, 434696, 434699), -946, -947 (gdb 434706, 434709), -1589-1591 (gdb 1238855-1238857), -1624, -1628, -1664 (gdb 3750105, 3750146, 3750161), ADU breakpoint (Augusseau *et al.* 1986; S579494)) (figures 24, 25A). In disagreement with earlier reports (Morrow *et al.* 1995), I map markers D22S946 and 947 to 22q11 contig 1 (figures 24, 25A) and marker D22S553 (Hudson *et al.* 1994, Budarf *et al.* 1996) to contig 3 (figures 24, 25C). Other markers and breakpoints, without a published sequence, are broadly mapped on the basis of available data (patient G (Levy *et al.* 1995), GM00980 (Fu *et al.* 1976) and GM05878 (Cannizzaro *et al.* 1985) breakpoints, markers D22S1618, -1637 and -66) (figure 24). Through FISH with fosmids overlapping with a contig 1-2 gap-spanning fosmid (f39g9), we find that the centromeric breakpoint of patient G is located on this clone (figure 24). FISH experiments show that cosmid 38a6 locates proximal to the GM05878 breakpoint (Mulder *et al.* 1995). McKie *et al.* describe the localization of *HUDDL/PNUTLI* to a point distal of the GM05878 breakpoint (McKie *et al.* 1997), which means that the breakpoint is positioned in an approximately 50 kb band between cosmid 38a6 and *HUDDL/PNUTLI* (figures 24, 25C, D). The GM00980 breakpoint is located in the approximately 80 kb range between cosmids 59f6 and 38a6 as indicated by FISH experiments (not shown and Mulder *et al.* 1995) (figures 24, 25C). Markers D22S1618 and D22S1637 represent the ends of cosmid 53e5 (B Morrow: personal communications) (figures 25C, D).

Recently, I re-analyzed the genomic sequence of the 22q11.2 region at the Sanger Centre using a sophisticated and highly automated analysis protocol geared towards dealing with large genomic sequences. Contigs 1 through 3 are now contiguous; the sequence of 236 kb stretching from contig1 to contig3 was briefly available under number U30597 (currently the sequence is only avail-

able as a series of separate cosmid entries). Figures 26 and 27 show some of the features of the sequence as displayed by the Fmap function of the ACeDB database and display program. Any new data from this re-analysis is discussed in the relevant paragraphs that follow. For the purpose of readability and compatibility with the first analysis the first three contigs will still be discussed as separate units. Figure 26 shows the relationship between the separate contigs 1-3 and the unified contig.

VII 1 22Q11 CONTIG 1

VII 1 1 ESTABLISHED TRANSCRIPTION UNITS

I mapped no new major transcription units to contigs 1 and 2. The genes positioned here, and their linear order along the contigs, have been reported by others (Budarf *et al.* 1995, Demczuk *et al.* 1995a, Wadey *et al.* 1995, Gong *et al.* 1996, Lindsay *et al.* 1996, Sutherland *et al.* 1996, Holmes *et al.* 1997).

The *DGCR5* gene, a gene without obvious protein encoding potential that can be transcript spliced into at least six different mRNAs (Sutherland *et al.* 1996), is present almost completely in the sequence of contig 1 (figure 25A). I can not locate the first 104 bp in its proper orientation, but a sequence 99% identical to this stretch is present in reverse orientation 5' of the gene. The fact that *DGCR5* is bound by an X-grail predicted CpG island 5' and polyA signal 3' indicates that the genomic range of the gene as shown in figures 25A and 27A is indeed complete. In the last intron (number 5) a 2300 bp single exon transcription unit is present, *DGSA*, that has homologous ESTs (including some with a polyA tail) and shares very high homology to a human membrane protein. There is contradictory information about the transcriptional direction. ESTs with polyA tails and the database sequence entry suggest forward transcription (centromere to telomere). However, the very high homology to proteins Q13116 and P52591 and cDNA U21556 suggest otherwise. For that reason it is indicated as a forward gene in figures 24 and 25A and a reverse pseudogene in figures 26 and 27A. *DGSA*, and *DDT2*, are different from other intronic and minor intergenic

transcripts in that more than one EST map to them and/or that at least some ESTs contain a polyA tail. *DGCR3* is a 260 aa ORF on the same strand as *DGCR5* that locates to the same *DGCR5* intron as *DGSA*. No ESTs or cDNAs map to *DGCR3* and no conserved sequences are found by database searches. Another potential transcription unit without cDNAs and ESTs, *DGSB*, has been mapped to the last intron of *DGCR5*. The current status of these potential transcription units is unknown.

Distal to *DGCR5* is *IDD*. The *IDD* cDNA sequence used for the alignment has a direct duplication of eighteen basepairs approximately 450 bp from the 3' end. The eighteen basepairs are present twice head to tail in the cDNA, but only once in the genomic sequences. *IDD* encodes a potential integral membrane adhesion receptor. Considering this functional classification, it will be of interest to determine the expression pattern of mRNA in mouse embryos. Until now, no mutations have been reported in non-deleted patients.

The fact that a mouse BAC of 130 kb (AC000096), carrying the 3' end of *Idd* at one of its ends (not shown), does not contain any *DGCR5* like sequence indicates that close to the 3' end of *IDD* the synteny between human 22q11 and mouse 16 is lost or interrupted. Interestingly, we do find *ZNF74* -like sequences downstream of the murine *Idd* gene; in *Homo sapiens* the *ZNF74* gene is located at the distal end of the commonly deleted region. On our 22q11 contigs I do not find a human homolog of this *ZNF74* like sequence.

D*GSG* and *DGSH*, both intronless transcription units, are situated telomeric of *IDD*. The 3' ends of the units overlap sixteen basepairs. *DGSG* is a serine threonine protein kinase. In the syntenic region of the mouse genome, apparently two functional copies of this gene are present (Galilli *et al.* 1997), both single exon genes (*Tsk1* (*Tssk1*) and *Tsk2* (*Tssk2*)). In human genomic sequence the ortholog of *Tsk1* is present as a pseudogene. There are as of yet no indications that the mouse has a *DGSH* homolog.

The fairly compact *DGSI* gene is positioned less than 1 kb distal to *DGSH*. One EST (C17097) suggests alternative splicing (addition of at least 108 bp to the 5' of exon 9) (figures 25A, 27B). The genomic

organization of *DGSI* was recently reported by Gong *et al.* (1997) and is in line with our findings. Gong *et al.* give no indication of alternatively spliced messengers. If the alternative splicing event can be confirmed experimentally, mutation analysis could be performed on the expanded exon 9 to complement the analysis already performed. The extra sequence added by this alternative splicing is much too short to be responsible for the 5.2 kb transcript (three times the size of *DGSI*) to which *DGSI* probes hybridize on Northern blots (Gong *et al.* 1997). *DGSI*'s function is not known but it is known to be highly conserved (homolog in organisms as remote as *C. elegans*) and it is known that the protein is located in the nucleus and that in mouse embryos it is highly expressed (Lindsay *et al.* 1998).

The sequence bridging the gap between contig 1 and contig 2 is approximately 30 kb in size. It contains the sequence for the *GSCL* gene (Galilli *et al.* 1997) (figure 27B). Like in the mouse genome this gene is positioned just distal from *DGSI*. Recently, Gottlieb *et al.* (1997) published their initial research on the human *GSCL*: expression is localized and at a low level. Adult pituitary and testis tissues are positive, as well as 9-10 week fetal thorax. Wakamiya *et al.* (1998) and Saint-Jore *et al.* (1998) mention the expression in the pons region of the developing brain and adult expression in brain, eye, thymus and thyroid. As this coincides with at least part of the CATCH22 phenotypes and the *Gooseoid* homeobox gene to which it is homologous plays a role in proper craniofacial development (Yamada *et al.* 1997), it was suggested that this gene is an excellent candidate gene. Experiments with *Gscl* null or hemizygous mice suggest that at the most it has a contributing role as these mice appear normal (Wakamiya *et al.* 1998, Saint-Jore *et al.* 1998). Mutational analysis in non-deleted VCFS patients did not reveal mutations in *GSCL* (Funke *et al.* 1997). One interesting fact is that *GSCL* and neighbor *DGSI* have some kind of relationship. Not only is *DGSI* expressed above average in the very pons region where *GSCL* is expressed, but in *Gscl* null mice, this expression domain is absent (Wakamiya *et al.* 1998).

VII i 2 NOTABLE FEATURES

When comparing the sequence of repeat-masked 40 kb fragments of contig 1 to our database of the five contigs, I noticed the presence of two stretches of approximately 900 bp that have 89% nucleotide homology and map 20 kb apart in identical orientations within contig 1 (30729-31636 on 103a2 and 17250-18155 on f41c7). The distal sequence is contained in *DGSA* (gray/black flags in figure 25A). This is the only occurrence of such an inverted sequence on the entire 1.2 Mb of 22q11 sequence. Database homology searches (EMBL/Genbank, human repeat and EST databases) fail to identify homologous sequences other than *DGSA*.

VII 3 22Q11 CONTIG 3

VII 3 1 ESTABLISHED TRANSCRIPTION UNITS

The three 5'-most exons of *CLTD*, to which no ESTs map, are found on contig 3 (figure 25C). We are unable to locate the first 53 bp of the *CLTD* cDNA anywhere in our 22q11 database. This stretch is not present in the other Clathrin-like clones (*CLH-22* (X95486) and *CLTCLI* (U60802)). Database comparison shows that this sequence is 98% identical (one mismatch) to sequences in the 5' ends of cDNAs of *MKK6* (Map Kinase Kinase 6 (U39656, U39657)), so I assume that the *CLTD* cDNA is chimeric.

Equivalents of alternative mouse Hira exons 3A and 11A (Wilming *et al.* 1997; X99713, X99722) are present in the sequence of the human *HIRA* locus (figures 25C, 27D), but amongst the few ESTs mapping to *HIRA* we do not find ESTs matching these exons. The possible existence of an additional alternative exon is indicated by cDNA clone zh90b02, which has a 3' end with homology to intronic sequences (short exon bar between exon 21 and 22 in figure 25C), whereas the 5' end contains sequences mapping to canonical exons (exons 16, 17 and 18). The genomic organization of the *HIRA* gene has been reported by Lamour *et al.* (1995) and is confirmed in our study. *Hira* is expressed in mouse and chicken embryos in structures affected in human DiGeorge patients (Roberts *et al.* 1997, Wilming *et al.* 1997, this thesis) and the gene product is likely involved in some aspect of maintenance, formation or regulation of chromatin structure. Possibly, *HIRA* is involved in opening or closing of chromatin of crucial developmental control genes and as such it is a candidate gene. For a detailed discussion see later in this section. Cosmids isolated with *HIRA* probes (table 4) are shown in figure 25C as gray bars (indicating that they are not sequence mapped but fingerprint mapped).

The recently described *UFDIL* gene is located 25 kb telomeric of *HIRA* (figures 25C, 27E). X-grail predicts a putative promoter with a CpG island in close vicinity. The published human *UFDIL* cDNA contains a sequence stretch duplicated into two consecutive copies, although some human transcripts are reported to lack the duplication of the stretch coding for 36 amino acids (Pizzuti *et al.* 1997). I fail to identify the corresponding duplication in the genomic DNA. The mouse cDNA sequence (U64445) is colinear with the human genomic sequence since it also lacks the duplication (Pizzuti *et al.* 1997). Five out of the six ESTs we map to the 5' end do contain the duplication, so it is unlikely that it is a cloning or sequencing artifact in the cDNAs. It is either an artifact in the genomic clone (the duplication could have been removed by recombination) or a polymorphism. *UFDIL* (which, incidentally, stands for Ubiquitin Fusion Degradation 1 Like) exons have been trapped from chromosome 22 cosmids by Trofatter *et al.* in 1995 (Trofatter *et al.* 1995). Our analysis reveals the intron-exon structure. Locus D22S183, which we mapped to the DiGeorge deletion region before, is part of intron 3. *UFDIL* is probably involved in protein degradation through ubiquitination

(Johnson *et al.* 1995). Accumulation of ubiquitin conjugates in migrating neural crest cells (Wunsch and Haas 1995) suggests that haploinsufficiency of *UFDIL* could indeed contribute to part of the *CATCH22* malformations. More on this in chapter IX-2-5.

VII 3 2 NEW TRANSCRIPTION UNITS

Between *HIRA* and *UFDIL* we find a new gene of at least four exons (figure 27D). This gene, *DDT9*, is constructed from sixteen ESTs. These ESTs are from cDNAs isolated from fetal brain, fetal heart, fetal liver and spleen, multiple sclerosis, senescent fibroblast and pregnant uterus cDNA libraries. There is a CpG island 5' of the gene, as well as two potential polyA signals at the 3' terminus. X-grail supports the exons and predicts more exons up- and downstream (figure 25C). Neither predicted nor mapped exons match homologous proteins or protein domains.

Downstream from *DDT9* another potential new gene *DDT34* was discovered upon reanalysis (figure 27D). The reverse gene matches rabbit ESTs C83394 and C82538.

Immediately telomeric to *UFDIL* I identify another new gene, *DDT12*, consisting of at least four exons and constructed from the 5' EST T34235 from white blood cells. There is partial support from X-grail, as well as an extra predicted exon (figure 25C). A CpG island is found upstream of the gene. The nearest forward polyA signals are located several kilobases downstream from the last mapped exon. Translations of potential *DDT12* coding regions were compared to protein and translated databases, but no significant homologies are evident. Continuing in the telomeric direction from *DDT12* I find a new gene we name *DDT14*, derived from homology to a mouse macrophage and a mouse E13.5-14.5 EST (AA071728 and W76825, respectively). With these ESTs I can construct a gene with a minimum of six exons, partially supported by X-grail. A recent study revealed that *DDT12* and *DDT14* are actually part of one transcript: *CDC45L* (McKie *et al.* 1998). *CDC45L* is, as the name suggest, homologous to *CDC45p*, a yeast cell cycle control protein. In retrospect the Grail predictions proved to be very accurate in that all but one of the *DDT14* predicted exons are indeed part of

CDC45L (compare figures 25C and 27E).

A notable feature close to, or in, *DDT14* is the occurrence of a 430 bp sequence duplicated (with 95% identity) on opposite strands 4 kb apart (black flags in figure 25C) (9686-10111 and 13980-14410 on 129f8). Database comparison with human repeat and EMBL/Genbank databases does not reveal any significant homology. I find a similar duplication on opposite strands between *CLTD* and *HIRA* (figure 25C). This 435 bp sequence (53914-54344 and 54820-55254 on 72f8) encompasses 280 bp of Alu sequence and is duplicated with 98% fidelity in two copies on opposite strands 480 bp apart. Database searches with the region outside the Alu homology do not result in the identification of homologous sequences. The duplication near *DDT14* is positioned in our cosmid contig surrounding marker D22S183. As described in an earlier chapter, with probe NB84 for this marker we isolated cosmids from the Lawrence Livermore chromosome 22 specific cosmid library (Mulder et al. 1995) and did a cosmid walk (figure 25C). FISH with two cosmids containing one or two copies of the duplication (97d9, 129f8) results in two signals per cosmid per chromosome 22 (in q11 plus pericentric) (not shown).

By comparison with translated EMBL nucleotide sequences I find sequences highly homologous to the rat androgen withdrawal and apoptosis induced protein *RVP1* (Briehl et al. 1991; Q63400) and mouse oligodendrocyte specific factor *Osp* (Bronstein et al. 1995, Bronstein et al. 1996; Q60771). I will refer to this gene as TRAINMODEL (similar To Rat Androgen withdrawal and apoptosis INduced protein, Mouse-Oligo-DEndrocyte-specific-protein-Like) in the rest of this thesis (figures 25C, 27E). X-grail predicts an excellent exon in part of the locus and another excellent exon upstream. On the nucleotide level there is no homology between *Osp* (U19582) and TRAINMODEL, but on the protein level identity is 33% over almost the entire length of *Osp*. The coding region is not interrupted by introns. Rat *RVP1* cDNA (PIR: A39484) is 65% identical to the TRAINMODEL locus, but only within the ORF. The protein coded for by the *RVP1* ORF is 43% identical to TRAINMODEL over almost its entire length. I find no ESTs for TRAIN-

MODEL and if the expression of rat *RVP1* mRNA is anything to go by, this is not surprising: *RVP1* is very weakly expressed in ventral prostate epithelial cells and in the epididymis (Briehl et al. 1991). The mouse *Osp* gene is located on chromosome 3, the human *OSP* on the syntenic human chromosome 3q26.2-q26.3 (Bronstein et al. 1995). *Osp* is only found in CNS oligodendrocytes and may be a homolog of PMP-22, a protein that plays an important role in CNS myelogenesis (Suter and Snipes 1995, Magyar et al. 1996). The diverging expression patterns and apparent functions of *Osp* and *RVP1* proteins do not provide clear cues about a possible function of TRAINMODEL. There is one long ORF coding for a protein of maximally 345 aa. The fact that several ESTs map to *DDT15* (indicative of a 3' end), that some of the ESTs contain a polyA tail, and that *DDT15* is transcribed in the same direction and positioned very close (79 bp) to TRAINMODEL, strongly suggests that *DDT15* is part of TRAINMODEL. Sirotkin et al. recently published a paper on the gene, which they call *TMVCF* (for Trans Membrane protein gene from VCFS region) (Sirotkin et al. 1997a) (the approved symbol is currently *CLDN5* for Claudin 5). They confirm the impression that TRAINMODEL is a single-exon gene. Their cDNAs suggest that a 219 aa protein is coded for. Computer analysis of the deduced protein sequence suggests that *TMVCF* is a transmembrane protein with two to four transmembrane domains. The homology with *OSP1* goes unnoted. In mice the gene is embryonically expressed from at least day 9PC, the earliest stage they tested. In man they detected transcripts in relative abundance in the adult heart, lung and skeletal muscle. Other tissues (brain, placenta, liver, kidney, pancreas) do express, but at a much lower level.

Restriction enzyme mapping and partial sequencing of several cosmids from a 22q11 cosmid contig we constructed earlier (Mulder et al. 1995) shows that cosmid 31e on the distal end of contig 3 is identical to our cosmid 31e3 (figure 25C). The sequenced contigs 3 and 4 do not overlap, but our cosmid contig does bridge the gap between them: we establish that cosmid 83h10 from the telomeric end of our 31e3 cosmid contig is virtually identical to clone carlaa from the centromeric end of our 22q11 contig 4 and that 100h is our 100h12 (figure 25D). Based on restriction enzyme mapping we estimate the distance between contig 3 and contig 4 at around 5 kb. In figure 25C I show that YAC y30h10 (Halford et al. 1993b) contains almost all of contig 3 and that it extends in the gap between contig 3 and 4. Cosmid 60c11 contains the end of the left arm of the YAC and cosmid 19d3 contains the end of the right arm. PAC b20-280 spans the contig 3-4 gap: it was isolated with a probe derived from cosmid 60c11 and spans the region from this cosmid to cosmid 91c8.

Telomeric of TRAINMODEL we observe a notable lack of EST and peptide homology over at least 180 kb (figures 25C, 26).

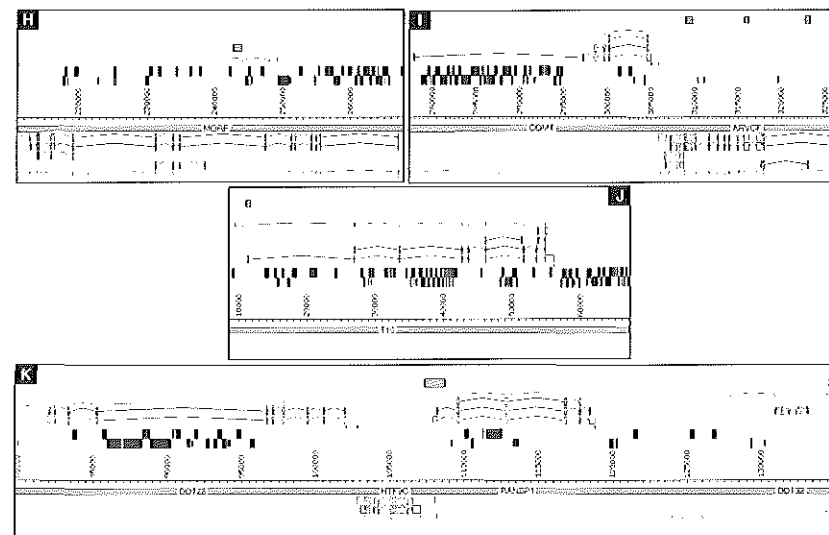
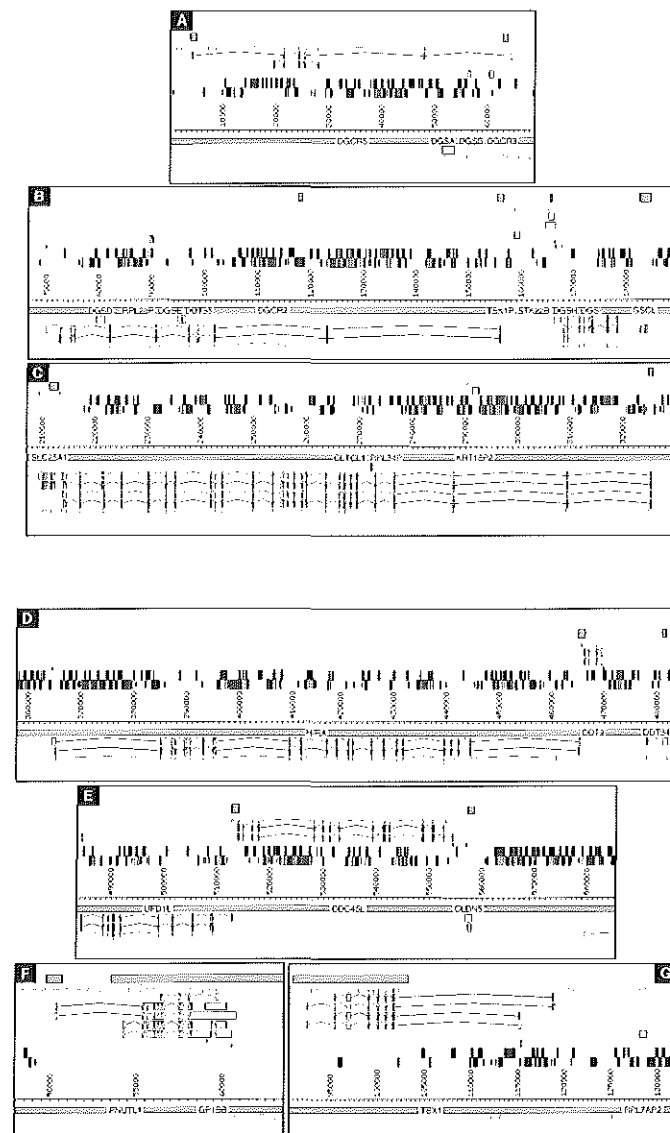


figure 27 (this is a grayscale thumbnail of the original for quick reference; see full size color picture on the foldout poster on the back cover)

Detailed view of all genes identified and discussed here. Imap outputs at varying scales show the intron-exon structure of genes (red: mRNA; green: protein coding sequence; blue: pseudo), the structure of predicted genes (narrow, light blue: Genscan; narrow, light red: Egenes), repeats (blue: long or LINE (LINE1 et al.); green: short or SINE (Alu et al.)), polyA sites and signals (narrow red vertical bars above repeats) and putative CpG islands (solid yellow boxes). Coordinates are as in figure 26. Genes below the yellow horizontal bar are reverse genes, above are forward (centromere to telomere) genes. Remaining features (repeats, CpG islands, polyA sites and signals) are not shown in a strand sensitive fashion and are always shown in the forward section. Where multiple gene objects are shown in parallel there are indications of alternative splicing. Some of these splice forms are partial, based on incomplete data (a single or few ESTs). Most alternative splice forms are not experimentally confirmed and some may be artifactual. **A:** DGCRC5 with, in the last intron, DGS5A, DGSB and DGCRC3 pseudogenes. **B:** DGCRC2 with pseudogenes/markers DGSB, RPL28P (ribosomal protein pseudogene), DGSE and DDT35 in its introns. Serine threonine protein kinase pseudogene TSK1P, overlapping genes serine threonine protein kinase gene STK22B and DGSB, and homeobox gene GSCL are also shown. **C:** Citrate transporter SCL25A1 preceding the large clathrin family protein gene CLIC1 and intragenic ribosomal protein and keratin pseudogenes RPL34P and KRTH2P2. **D:** Chromosomal/transcriptional regulator protein gene HIRA and DDT19 and DDT34. **E:** Ubiquitin pathway gene UFDL1, cell cycle control gene CDC45L and possible transmembrane protein gene CLDNS. **F:** Homolog of cytokinesis genes PNUTL1 and platelet glycoprotein GPIBB. **G:** T-box gene TBX1 and ribosomal protein pseudogene RPL2AP2. **H:** Reductase family member MORE1. **I:** Neuronal gene COV1 and catenin family member and possible signaling protein gene ARVCF. **J:** T10, ortholog of the mouse gene with the same name. **K:** DDT28, possible RNA interacting protein gene HTF9C, nuclear regulator RANBP1 and DDT32.

bp sequence stretch with a very high adenyl content, suggestive of the remnants of a polyA tail; and fourthly, *RPL7AP2* is not clustered like *SURF3* in the highly conserved *Surfeit* genomic cluster of apparently unrelated genes. This arrangement is conserved across species (Williams *et al.* 1988; Colombo *et al.* 1992), suggesting important regulatory reasons. In the mouse pseudogenes of *surfeit* cluster genes have been found outside the cluster (Huxley *et al.* 1988). I think *RPL7AP2* is a processed pseudogene of *SURF3* which may be transcribed. In contrast to most other pseudogenes the protein coding potential of *RPL7AP2* is still intact. Since the original gene with identical protein coding potential is present elsewhere in the genome and it is a question whether *RPL7AP2* is functional to begin with, hemizygoty of *RPL7AP2* is very unlikely to have any effect.

A cluster of transcription units is located telomeric of *RPL7AP2* (*DDT17-20*) (figure 25D). No ESTs connect any of the separate units, so the relationship to each other is not known at this point. The first unit has a polyA tail (and a polyA signal) and several kilobases upstream of the fourth unit a CpG island is identified. Together with their close vicinity and the fact that all units are transcribed in the same direction, this suggests that this constitutes one gene. However, X-grail does not support any of the mapped EST matches.

The next transcription unit I identify is *MORF* (Member Of the Reductase Family), a gene coding for a protein homologous to *C. elegans* Glutathion Reductase (*Glured* (P30635)) and human Thioredoxin Reductase (*TXNRD1* (Q16881)) (figure 25D). Two of the mapped exons, the ninth and tenth, are subject to alternative splicing according to matches with ESTs (T66908, T66909, T67208, T67209). X-grail supports the vast majority of mapped exons. I recently mapped an additional potential alternatively spliced product that uses an extra exon in intron 6 (figure 27H). Alignment of translated *MORF* genomic DNA to human (P00390) or *C. elegans* *GLURED*, shows that *MORF* is more homologous to the *C. elegans* protein than to the human protein. Homology with human *TXNRD1* is on the same level as the homology with *C. Elegans* *Glured* and various other types of reductases also show peptide homology to *MORF*. Trofatter *et al.* mention two Blastx homologies

of exon-trapped chromosome 22 sequences to two ESTs with similarity to *Glured* (Trofatter *et al.* 1995). *GLURED* (which gene maps to human chromosome 8p21-p23) is involved in protecting cells from oxidative stress. At the moment, I do not see an obvious role of haploinsufficiency of this gene in *CATCH22* etiology.

A human homolog of mouse *p120/Catns* (Reynolds *et al.* 1994,1996; Z17804, P30999) maps less than 2kb telomeric of *COMT* (figure 25D). The gene will be referred to as *TOMCAT* (similar TO Mouse CATns). A human homolog of *Catns*, *CTNND*, was cloned recently (Reynolds *et al.* 1996). Comparison of mouse *Catns* peptide sequence with translated genomic sequences from the *TOMCAT* locus indicates an overall identity/similarity of 58%/82%. On the nucleotide level homology is weak, with homology between only a few exons (identity in the 56-65% range). The fifth exon is present in one cDNA (ESTs H17198, H17975), whereas in another clone (ESTs T79735, T79650) the fourth exon is spliced to the sixth exon, skipping number five. This probably indicates alternative splicing. Recent reanalysis gave another EST (AA779605) indicating alternative splicing, this time with an exon in intron 1 (figure 27I). A CpG island several kb 5' of the gene could indicate the presence of a promoter region for *TOMCAT*.

Mouse *p120/Catns* is a member of the Catenin family. The human *p120* homolog *CTNND* has been assigned to chromosome 11q11 and the mouse gene to the syntenic chromosome 2. Mouse has four isoforms of *Catns* (*Cas1A*, *Cas1B*, *Cas2A*, *Cas2B*) two of which — *Cas1A* and *B* — are shown to be coded for by one alternatively spliced gene (Reynolds *et al.* 1994). Interestingly, the evidence we find for alternative splicing in the 3' region of *TOMCAT*, is approximately 100 bp from the equivalent position in *Catns* (Reynolds *et al.* 1994). *TOMCAT* lacks homologous sequences for two stretches of N-terminal *Catns* amino acids, similar to *Cas2A* and *B* *p120* isoforms; these isoforms (from (an) as of yet unidentified gene(s)) lack N-terminal sequences relative to the two long forms *Cas1A* and *B* (Reynolds *et al.* 1994). At present, all indications point to *TOMCAT* being the equivalent of the mouse gene coding for *Cas2A* and *Cas2B* (which gene may map to chromosome 16) and *CTNND* being the true equivalent of *Catns* (*Cas1A* and *Cas1B*). A recent publication by Sirotkin *et al.* (1997b) describes the isolation and analysis of the *TOMCAT* gene, there called *ARVCF* (for Armadillo Repeat protein gene from VCFS region). Mention is made of alternative splicing in the 5' end of the gene. While alternative splicing in the 3' end is not alluded to by Sirotkin *et al.*, they do note that the 3' ends of *ARVCF* and its proximal neighbor *COMT* overlap. I, on the other hand, find *TOMCAT* polyA tails at a position well distal (almost 1kb) from the *COMT* polyA tail. The clone which Sirotkin *et al.* refer to as overlapping with *ARVCF* is actually one of the constituents of *DDT23* (figure 25D). This is only one cDNA clone out of the 26 canonical ones, six of which have a polyA tail (all at the identical position). If the extended clone is not an anomaly, then at least it is not the canonical transcript. I have my reservations about this transcript, seeing that in my study *HIRA*, *T10* and *DDT32* also have anomalous transcripts mapping to their 3'

ends (*DDT6*, *DDT26* and *DDT33*, respectively). Until further experimental evidence shows that this cDNA (ESTs F08001, Z41079) represents part of a real transcript, for all practical purposes I consider *COMT* and *TOMCAT* genes as not overlapping.

Sirotkin *et al.*'s expression studies show that *TOMCAT/ARVCF* is ubiquitously expressed in the adult and embryonic human. Low stringency hybridization to genomic DNA from other species indicate that *TOMCAT* is conserved in vertebrates and *Drosophila*. The protein tyrosine kinase substrates CTNND and Catns are thought to be involved in ligand induced signaling by EGF, PDGF and CSF1 receptors. PI20 proteins have domains in common with *Drosophila* armadillo, human plakoglobin and *Xenopus* β -catenin (Reynolds *et al.* 1992). These molecules play a role in coupling the cell adhesion molecule cadherin to the actin element of the cytoskeleton. PI20 is associated with complexes containing E-cadherin, α - and β -catenin and plakoglobin (Reynolds *et al.* 1994). Factors involved in cell movement or adhesion are likely to play a role in developmental processes involving the highly mobile neural crest cells. *TOMCAT*, like *IDD*, might be such a factor.

Four ESTs (AA015875, AA015620, AA017589, AA017356) map to introns of *TOMCAT* (*DDT24* (figure 25D)). These are likely to be derived from genomic contamination.

VII 5 22Q11 CONTIG 5

VII 5 1 ESTABLISHED TRANSCRIPTION UNITS

Mouse *T10* (mapping to mouse chromosome 16) has been mapped in the human DiGeorge deletion earlier (Halford *et al.* 1993c). We find this gene present on 22q11 contig 5 (figure 25E). The X-grail predicted promoter in conjunction with a CpG island makes a good candidate for the promoter region of *T10*. Alternatively spliced mRNAs were later found through matches with ESTs AA459085 and H55355 (an exon trapped clone). The product matching AA459085 concerns an extension of an exon, which is supported by FGENES and GENSCAN gene predictions. (figure 27J)

AYAC of approximately 480 kb containing *T10* was isolated with probe HP500 (Halford *et al.* 1993c) for marker D22S932. Further analysis shows that this YAC also contains the *COMT* gene and that it thus spans the gap between contigs 4 and 5 (figures 25D, E).

VII 5 2 NEW TRANSCRIPTION UNITS

Amouse EST from an E13.5-14.5 embryonic cDNA library maps to 22q11 contig 5 in five separate exons, with a nucleotide homology of 93% (*DDT28* (figure 25E)). No human ESTs map here and no gene features are apparent in the vicinity of this gene, except maybe a CpG island approximately 5 kb upstream. All mapped exons are supported by Grail and in close vicinity it suggests one more 5' exon and three more 3' exons. I am unable to find proteins homologous to the potential trans-

lations of the *DDT28* exons.

Several ESTs (*DDT30*) map to a region downstream of *DDT28* and they are transcribed in the same direction (figure 25E). *DDT30* maps very close to the 3'-most predicted exon but does not overlap. New data available after initial analysis shows that *DDT28* is larger than originally found and that it contains all but the first predicted Grail exons (figure 27K). *DDT30* is not part of the gene.

Two ESTs map less than 1kb telomeric of the above mentioned groups, yielding four exons of *DDT31* (figure 25E), all Grail supported; Grail also predicts eight more exons 5'. Database comparisons of translated *DDT31* sequences reveal no homologies. Detailed reanalysis shows however that this gene codes in fact for the homolog of mouse Htf9c (Bressan *et al.* 1991, Guarguaglini *et al.* 1997; P70222). The true extent of the eleven exon gene includes *DDT31* and all Grail predicted exons except the most 3' (figures 25E, 27K). In mouse this gene maps to chromosome 16 and shares a bidirectional promoter with the *RANBP1* gene discussed next. Htf9c has some homology to RNA binding proteins and RNA methyltransferases.

Comparison of translated 22q11 contig sequences with translated yeast and *Drosophila* sequences led to the assignment of human *RANBP1* (Bischoff *et al.* 1995; D38076, P43487) to chromosome 22q11 contig 5 (figure 27K). X-grail predicts a CpG island and a promoter at the 5' end (figure 25E). Two cDNA clones give an indication of two different alternative splicing events (ESTs g90b09 and yz61g09). The ESTs mapping to this transcription unit are identical to human *RANBP1* cDNA (D38076, P43487) (100% protein identity). The sequence of the *RANBP1* gene was reported by Bischoff *et al.* (1995) and the name is short for RAN Binding Protein 1. *RANBP1* plays a role in regulating the RAS related nuclear protein RAN (Ras Associated Nuclear protein). RAN is part of a regulatory pathway implicated in regulating various nuclear processes: DNA replication, mitosis, mRNA processing and transport, protein import. *RANBP1* regulates RAN's GTP binding and hydrolysis and works in complex with RANGAP1 (RAN-specific GTPase-Activating Protein) and RCC1 (Regulator of Chromosome Condensation) (Bischoff *et al.*

1995). RANBP1 seems to be involved in a plethora of nuclear regulatory pathways and we can not exclude a role for the protein in CATCH22.

The last transcription unit on contig 5 is constructed from a mouse EST (AA107820) (*DDT32* (figure 25E)). The two mapped exons are supported by X-grail and ten more exons are predicted in a 4 kb region 5' of *DDT32*. When the potential ORFs are translated and compared to databases, I could not find significant homologies. Interestingly, more recent analysis indicates that the EST is not part of the gene located here. Its eight exons coincide with all except the last two Grail exons and match EST AA325327 (figure 27K). Some exons have homology to worm, yeast and plant predicted proteins and a human DHHC domain containing cysteine rich protein (O15461).

A group of four overlapping human ESTs overlaps with the second mapped exon of *DDT32* (*DDT33* (figure 25E)).

VII 5 3 NOTABLE FEATURES

When genomic clones 68a1 (AC000083) and 24b (AC000075) are compared to clone p201m18 (AC000097), two direct repeats are apparent that are not present on p201m18 but are on the other two clones. On 68a1 a 25 bp sequence is present in two consecutive identical copies. p201m18 has only one copy of this 25 bp stretch. Cosmid 24b has two copies of a 12 bp sequence that flank a 305 bp region. Again, clone p201m18 lacks the 305 bp region and one of the two 12 bp stretches. For the construction of contig 5 we used the longest sequences, i.e. from 68a1 and 24b, where there was disagreement between clones.

VII 6 MINOR INTERGENIC AND INTRONIC TRANSCRIPTS

On all five 22q11 contigs I find a total of seventeen transcription units (*DDT1*, *DDT3-DDT8*, *DDT10*, *DDT11*, *DDT13*, *DDT16*, *DDT21*, *DDT23*, *DDT25-DDT27*, *DDT29*) that generally are constructed from only one cDNA and are intronless as far as I can determine.

They overlap with exons, or are located between genes or in introns of genes. Some show homology to high copy repeats. Their transcriptional direction is invariably the same as for the "host gene" when they are located within a gene. *DDT1* is an exception, but this may be a reversed clone.

VII 7 LOW COPY REPEATS

In the 1.2 Mb stretch of chromosome 22 sequence discussed here I found evidence of only two pairs of low copy repeats, with the elements of each pair relatively close together (figure 25C). The main impediments in our attempt to create a cosmid contig were gaps and spurious hybridizations. Could these be indications of low copy repeats? The failure to find overlapping cosmids after a short walk (started with probe NB84) is in line with reports that the 22q11 region is notoriously hard to clone. This could be caused by an inherent instability of the region, which in turn could be caused by low copy repeats. It has been suggested that in the past (on evolutionary scale) chromosome 22 was much larger than it is at present. Repeats scattered over the chromosome would have been responsible for numerous intra-chromosomal recombination and subsequent deletion events, weeding out non-essential sequences and reducing the chromosome to its current size. This ongoing process may be reflected in the observed instability. Only now it has fatal results as witnessed by CATCH22 and other syndromes. Halford et al. (1993a) described low-copy repeats in and around 22q11. The spurious hybridization and FISH results — i.e. strong cross-hybridization of non-contiguous cosmids and multiple signals on chromosome 22 and other chromosomes when FISHing with certain cosmids, despite masking of high copy repeats — can be attributed to either low-copy repeats or to rearrangements (which in turn may be attributable to low copy repeats). Whether the two repeats we see in the 1.2 Mb 22q11 sequence are in any way related to the repeats discussed above is currently unknown. The inverted repeats in contig 3 are located on a cosmid that gives two chromosome 22 signals when used in FISH experiments (cosmid 129f8).

VII 8 GENERAL DISCUSSION OF DETAILED ANALYSIS OF DGCR GENOMIC SEQUENCE

With the information obtained from the above discussed low resolution mapping and with the advent of HTGS (High Throughput Genomic Sequencing) projects, one of them directed towards the DGCR, I analyzed 1.2 Mb of submitted DGCR sequence in detail.

I set out to identify transcription units that are evolutionary conserved and homologous to developmental control genes, using biocomputing supplemented with experimental data. From the sequences in the Genbank/EMBL database, I constructed five non-overlapping 22q11 contigs and searched for expressed and conserved sequences by comparison with EST, yeast and fly sequence

In discussion's context not possible. These items are included with a name starting with DDT. Updated analysis available on the website of the National Center for Human Genome Research. All rights reserved. Copyright 1998 by Cold Spring Harbor Laboratory Press.

databases. Here I report the position and genomic structure of twenty-seven genes within these five contigs; of these, twelve were known and the existence or location of fifteen was unknown at the time of analysis.

Of the twenty-seven genes described in this chapter, eleven have homologous sequences in yeast, worm and/or fly (*GSCL*, *CLTCL1*, *HIRA*, *DGSI*, *UFD1L*, *HUDDLE* (*PNUTL1*), *RPL7AP2*, *MORF*, *DDT31* (*HTF9C*), *RANBP1*, *DDT32*). Except *GSCL* none of these homologous genes are obvious developmental control genes in the worm or the fly. *IDD*, *DGSG*, *CDC45L* (*DDT12* + *DDT14*), *TBX1*, *TRAINMODEL* (*CLDN5*), *TOMCAT* (*ARVCF*), *T10* and *COMT* are conserved in mammals. *TBX1* is the only potential developmental control gene in this selection.

My study shows that the X-grail program is very effective in confirming exons known to exist from homology to cDNAs, ESTs or proteins. Virtually all coding exons described in this thesis, around 170, are considered excellent by at least one X-grail version and mostly by all three X-grail versions. The predicted intron-exon borders on the genomic DNA (Grail 2) are usually identical or very close to the borders determined from alignment with cDNA. Only few exons are missed, which can be due to sequence errors obscuring splice sites or ORFs, or to the use of non-consensus splice sites. The high success rate of the program could have a cause in the way the program is trained. Grail is a neural network type program that is trained on existing genes and is subject to permanent education: as the structure of new genes becomes available, the program is fed these intron-exon structures, giving it the opportunity to refine its algorithms.

Grail predicted exons that turn out to be real are usually grouped, i.e. not interspersed with other 'excellent' predicted exons. Thus, a cluster of excellent exons appears a good indication of the presence of one or more genes (this probably doesn't work for large genes with large introns). Based on this premise, I give the possible genomic organization of *DDT9*, *DDT12*, *DDT14*, *TBX1*, *MORF*, *DDT28*, *DDT31* and *DDT32* (figures 25A-D). Since the initial mapping of these genes is based on possibly incomplete and/or imperfect data (protein and/or EST alignment), I might have missed many exons. With Grail predictions I extend the mapped portion of these genes with predicted exons, but nevertheless the genomic organization I show is most likely not complete nor completely accurate and experimental analysis will have to show the true and complete organization of these genes. For some of these putative genes experimental evidence has come to light after my initial analysis. In all of these cases the majority of predicted exons as annotated in figures 27A-D turn out to be confirmed (*DDT12* + *DDT14*, *TBX1*). Where reanalysis showed new and extended matches to recent ESTs, proteins and other sequences the same holds true (*DDT28*, *DDT31*, *DDT32*). This shows that the combination of predictions and up-to-date homology searches is a very powerful one.

X-grail predicted CpG islands are found upstream of twelve genes, including four in conjunction with a potential PolII promoter. Pending experimental confirmation, X-grail seems to be able to

predict promoter areas for more than 40% of the genes. Prediction of potential promoters alone is not sufficient as the rate of false positives is too high. The false positive rate for CpG islands seems to be much lower, so the presence of a predicted CpG island (alone or in conjunction with a predicted promoter) combined with a cluster of excellent exons, has proven to be a fairly good predictive tool.

I find at least seventeen intronless intronic and intergenic transcripts, which sometimes contain, or overlap with, exons of (established) genes. Invariably these transcripts are based on only one or two ESTs (one cDNA clone) and they are always transcribed in the same direction as the host gene when they are intronic. Some of the ESTs are homologous to Alu repeats. Transcripts *DGSB*, *-D*, *-E* and *-F* as described by Gong *et al.* (1996), fall in the same category and we think that these transcripts do not belong to conventional genes but are associated with high copy repeats or originate from cDNAs primed on repeat associated polyA stretches in genomic DNA contamination. Further experimental research should reveal more about the nature of these minor putative transcripts. We are sure they play no role in CATCH22 etiology.

The number of ESTs found for each gene varies widely. Obviously the number of hits is a function of the abundance and distribution of the mRNA. This probably explains why *TBX1* positive ESTs were not identified: *TBX1* expression is quite spatially restricted. The low number of hits for *CLTCL1* indicates low and/or localized expression, which is in keeping with the muscle specific expression described before (Sirotkin *et al.* 1996). Judging from the number of EST homologies, *SLC20A3*, *IDD*, *UFD1L*, *HUDDL*, *COMT*, *TOMCAT*, *DDT19* and *RANBP1* are highly and/or ubiquitously expressed.

With nine genes (*DGSI*, *HIRA*, *HUDDL*, *GPIBB*, *MORF*, *COMT*, *TOMCAT*, *T10* and *RANBP1*) I have reason to assume the occurrence of alternative splicing that has not been described before. In these instances detailed alignment of ESTs with genomic DNA reveals cDNA clones that contain an extra exon, lack (an) exon(s) and/or use alternative splice sites. Some of these EST based predictions could well reflect artifacts, but pending experimental confirmation, detailed EST analysis appears to be a very powerful tool for

detecting alternative splicing events.

In general the linear order of chromosomal markers is as described earlier (Budarf *et al.* 1996, Morrow *et al.* 1995), but markers D22S947 and 553 are located at different positions than previously described. According to Hudson *et al.* (1994) and Budarf *et al.* (1996) D22S553 is located in their bin 5/6, which is telomeric of our contig 4 (and possibly even beyond 5). I map the marker to contig 3 (located in bin 3), just downstream of HIRA. This is in keeping with the position reported by Morrow *et al.* (1995). Regarding markers D22S946 and 947 however, their results disagree with my findings: Morrow *et al.* indicate the two markers as locating distal to D22S553, 609, 942 and 941, 944 and 943 (which, we show, all map to contig 3), while I positively map the markers to contig 1, i.e. proximal to the mentioned markers. Marker D22S75 (McDermid *et al.* 1989) is located in the gap between contigs 1 and 2. Probe N25 for this locus hybridizes to *CTP* (Mulder *et al.* 1995), but comparing restriction maps of the area surrounding *CTP* with those of N25 (McDermid *et al.* 1989) we do not see any homology. Thus D22S75 must locate to the gap. Indeed, Gottlieb *et al.* (1997) recently showed that N25 hybridizes to the *GSLC* locus (we assume upstream of the gene), which is located in the gap in question (figure 26)

We have described the position of cosmids 38a6 and 33d9 relative to the GM00980 and GM05878 breakpoints above and in Mulder *et al.* (1995). In the detailed study discussed here I narrow the positions of the GM00980 and GM05878 breakpoints down to a region of 80 kb on contig 3 and 50 kb on contig 4, respectively. At the moment there are no indications that there are genes disrupted by these breakpoints, since they are located in a larger than 180 kb region of 22q11 where we fail to find indications for the existence of genes.

Sequence analysis was done in collaboration with the originators of the genomic sequence (Dr Emanuel's group in Philadelphia and Dr Roe's group in Oklahoma). Raw sequence data was made available for us to analyze. I had previously constructed various cosmid contigs (chapters VI-2 - VI-7), three of which overlap with the cosmids used

for the large scale sequencing. The M56 contig adds extra information to the sequenced contigs by bridging the gap between contig3 and 4. Recently, the sequence has become available (Dunham *et al.* 1999). We made our annotation of the sequence available to the Emanuel and Roe groups to aid in their search for and study of new potential candidate genes. Studies on several of the new genes are under way as a result. A joint report describing the results of experimental data from our collaborators combined with our analysis and annotation is currently in progress.

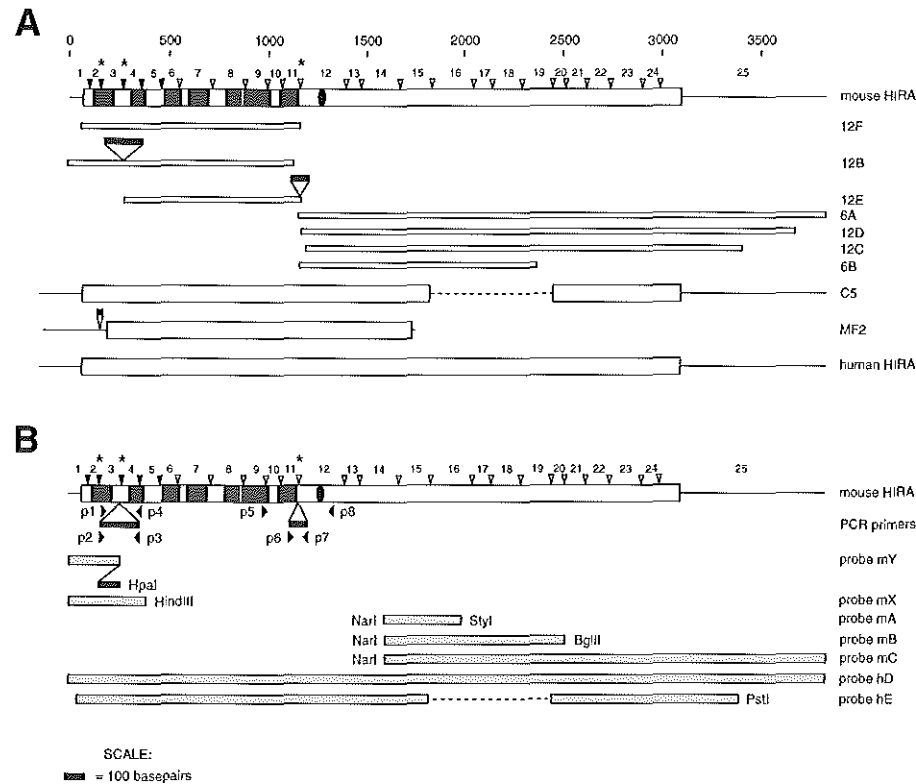


figure 28 Schematic drawings of HIRA cDNAs, cDNA clone inserts and probes. **A:** The assembled mouse cDNA and the individual cDNA clones isolated from an EH.5 mouse embryo cDNA library (12F, 12B, 12E, 6A, 12D, 12C and 6B). As a reference the human and murine Tup1^l C5 and MF2 cDNAs and human HIRA are shown. Nucleotide positions are indicated on top. Open boxes depict open reading frames; dark gray boxes depict WD40 domain coding regions; black oval indicates penta-Q stretch coding region; closed arrowheads mark exon-exon boundaries determined from mouse genomic sequences; open arrowheads mark human exon-exon boundaries (exon numbers are between the arrowheads); asterisks above arrowheads show the position of additional exons 2A, 3A and 11A; black bars indicate the additional exons. **B:** Drawing of a putative mouse Hira transcript with additional exons and the positions of the various primers and probes used. Probes preceded by m are mouse probes, the ones preceded by h are human probes. Restriction sites on mouse or human HIRA are given as a reference. from Wilming et al. (1997).

VIII ANALYSIS OF HIRA (DGCR1, TUPLE1), A CATCH22 CANDIDATE GENE

In the previous chapters I described the low-resolution and high-resolution mapping of cosmids, probes, markers and genes in the 22q11 region. *HIRA/TUPLE1* has been mentioned several times as a probe in the low-resolution mapping section and as a candidate gene in the high-resolution mapping section. The gene is localized in the vicinity of the ADU t(2;22) translocation breakpoint (Halford et al. 1993c), a breakpoint of special interest because it represents a fairly unique combination of a balanced translocation combined with a DiGeorge phenotype. This strongly suggests that this breakpoint disrupts coding or regulatory sequences essential for the function or transcription of a gene in its vicinity. Though *HIRA/TUPLE1* is located an estimated 150 kb from the breakpoint (Halford et al. 1993c), *trans* acting regulatory sequences could be located at the ADU breakpoint. Similar cases have been shown for *SOX9* (Wirth et al. 1996) and *PAX6* (Fantes et al. 1995). Original best homologies of *TUPLE1* were with yeast *Tup1p* and *Drosophila* Enhancer of Split (hence the name *TUP1* Like Enhancer of split). Later, best homology was with yeast *Hir1p* and *Hir2p*, resulting in the renaming of *TUPLE1* to *HIRA* (Lamour et al. 1995). Sequence analysis by Halford et al. (1993c) suggests that *HIRA/TUPLE1* possibly functions as a transcriptional regulator. This warranted a close study of *HIRA/TUPLE1* to determine its CATCH22 candidate gene status. To that effect we studied (embryonic) expression, concentrating on the mouse homolog.

VIII 1 HUMAN HIRA

A probe from the insert of human *TUPLE1* cDNA clone C5 was used to screen a 17-18 week human fetal brain cDNA library. We isolated three clones: HT4, HT5 and HT9. Restriction enzyme analysis revealed that, compared to the C5, the 4.5 kb HT4 contains

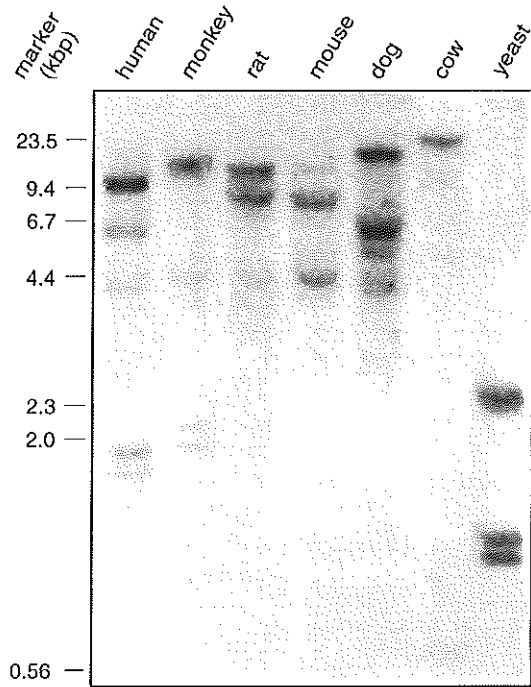


figure 29 Conservation of the HIRA gene. A human HIRA probe, hD, detects conserved sequences in *EcoRI* digests of genomic DNA from human, monkey, rat, mouse, dog, cow and yeast. From Wilming et al. (1997).

extra sequences inserted in the 5' region, as well as extra sequences in the 3' region. HT5 (4 kb) only has the 3' extra sequences but extends 200 bp further 3' than C5. HT9 (2 kb), like HT5, does not have the extra 5' sequences, but extends 150-250 bp further 5' than C5. Database searches with the end sequences of HT9, show that this 5' extension consists of ribosomal RNA sequences: HT9 is chimeric. HT4 and HT5 could represent splice variants. With the mouse as our experimental animal of choice, and to confirm the existence of splice variants, we were interested in obtaining similar mouse cDNA clones.

VIII 2 MURINE HIRA cDNA CLONES CODE FOR A PROTEIN WITH HOMOLGY TO HUMAN CHROMATIN ASSEMBLY FACTOR 1A AND MULTIPLE ALIGNMENT SUGGESTS THE PRESENCE OF THREE FUNCTIONAL REGIONS IN HIRA

We examined the evolutionary conservation of *HIRA* by hybridization of radiolabelled probe hD (from the 4 kb insert of one of our human *HIRA* cDNA clones) (figure 28B) to a blot containing *EcoRI* digested genomic DNA from various species. The probe hybridizes to homologous sequences in the DNA of all mentioned organisms (figure 29). This high level of conservation from man to yeast indicates an essential function for *HIRA*.

As an initial step to study whether haploinsufficiency of the *HIRA* gene contributes to the CATCH22 phenotype, we hybridized probe hE (figure 28B) to an E11.5 mouse embryo cDNA library under stringent conditions. This screening yielded seven positive cDNA clones of the mouse homolog of the human *HIRA* gene (figure 28A). Complete sequencing and assembly of selected clones allowed us to construct a cDNA of 3852 basepairs, containing an ORF of 3045 basepairs. The 1015 aa predicted protein has seven WD40 repeats (at aa positions 10-44, 68-98, 129-159, 172-202, 234-259, 266-313, 326-356) two bipartite nuclear localization signals (at aa 267-286 and 626-643), a penta-Q stretch (aa 408-412) and a molecular mass of 111.72 kD.

Sequences comparison between the human and mouse proteins reveal an overall identity of 95.8% and a similarity of 97.2%. The N-terminal one third of the protein, containing the WD40 repeats, is 99.4% conserved (identity and similarity), leaving the C-terminal two thirds with relatively higher divergence (identity 93.7%; similarity 96%). Two non-conserved amino acids are present in WD40 repeats 3 and 6. Compared to the 1017 aa human *HIRA* translated sequence, the mmHira sequence lacks amino acids 510 and 607 or 608. Our Hira sequence is 99% identical to the recently reported mouse Hira sequence (Scamps et al. 1996) and 100% identical to the mouse TUPLE1 cDNA clone MF2 (Halford et al. 1993a, b, c).

In addition to homology with yeast Hir1p and Hir2p, BlastP and BEAUTY database searches with our mouse Hira sequence reveal high homology to the p60 subunit of human Chromatin Assembly Factor I (CAF1A) (Kaufman et al. 1995) and a *Caenorhabditis elegans* protein (PID g687860) (which, in the absence of a name, we will henceforward refer to as *C. elegans* HIRA Like Protein (ceHILP). The overall homology with ceHILP is higher than with Hir1p (p-value 6.1×10^{-77} versus 7.6×10^{-45}). We aligned the murine Hira protein sequence with ceHILP, Hir1p, Hir2p and CAF1A (figure 30). The overall similarities in this alignment are 39% with ceHILP, 38% with CAF1A, 35% with Hir1p and 31% with Hir2p. It can be readily seen that there are at least three different conserved (functional?) regions: the highest homologies are found in the N-terminal part, region A (aa 1 - 371) (average similarity 47%); in the middle part, region B (aa 473 till aa 658), homology


```

hsCAF1A 372 E T E . X D E L G I P L K . E K P V N M R T P .
ceHILP 639 V H L S . S . V G C M S N . E A M S D L C Y Q I Y S I R P P R Y L T D K E E D E .
mmHIRA 354 E D E S . D E L G I P L S E E K S R . H S A Y G K S L A I . T A L S T A . E N P M E K . Q . R Q Q Q Q Q
yHIR2 348 T F A Q . K D . L G V A L P . T E I K S . Q . V N . L L P K L E E P L . Q I P K S F P E N I K L E .
yHIR1 382 L F R . . N . E I G K P . P L . E R N M E Q L Y R . G V D K D S . D F F . S I N Q . L E D . T K S . K . T K I S T S

hsCAF1A 396 . . . . . T A . K T K S Q T H R G S S P G P R . T E G . P A S R T
ceHILP 681 . . . . . S Q D . F . L S D L S . S A N N A S F V C
mmHIRA 413 L D Q K N A T T R E N S S A S S V T G V V N G E S L E D I R K L L K Q V E T R . A . G R R I . E . C I A Q L D T G
yHIR2 398 . . . . . E S . S A A P I P N D T G
yHIR1 440 K L G E N H P P L A N S A S . . . . . N Q . . . . . D N N D A S V S R S E H

hsCAF1A 425 Q D P . P G T . P Q . . . . . R Q A P A T V I R D P S I P A V K P L P P S E E T L Q P S
ceHILP 703 P E D V L I K R K K L . Q Q S S D T Q L K S M . D N S K E . E S K M S E K . M M E E R N .
mmHIRA 473 D F . . F F N . I P . . S L A G T . I S . P S G Q . L E . . S A S F S G . S K P C T E P . . S A P T G E S
yHIR2 411 R S . V . K K P . K K K . N N Q T N G . K . I Q S T S . E F N T P S Y T V P R D L K R K P K E A . P E N I A P G S K K
yHIR1 470 I N I L P K R K K D A I L N K A V . T K S G K K R V A P T L S . S S S S P F N G I K K P T . D . K R I E N N V K S

hsCAF1A 473 S Q N T K A H P R R V . L N . L Q A W S K . . . . . R I N T P L K . D T P P S V P F S V I S . P S . E E I Q S E T
ceHILP 750 K Q I . V R K D G K R R I Q P V F C G . . T . A A D P . T S . S S E G R . T V A P P P K R K K L . P . V P A K K C E V
mmHIRA 533 V . E . S M N A . . T P . . S P S V L . . . . . P I S . E . P . K . F D . . E R S K . I P G A P S . L . . I P T . V E
yHIR2 471 K K L Q P I D F L D . L L L P N T S F . R I R L A T . K I R S T F . . . . . P I N N E N L I L D V N G S G N E Q R P
yHIR1 530 S . N . T I N S K N . L L N V P E G V E K K I . S I S F P L . R L . I E . L I M G T K E R . S A W K I S N . E . E N D D A D

hsCAF1A 532 P G D A . G S P P P . E . R P R L . E N K G G T E S . D P .
ceHILP 808 D P E . S S D S D . D D E E . E E E D M E I . S D I E S V R N . K R R K . P A T T S . G P M . A L D L K R P A L R E M E R K S M
mmHIRA 591 R . K . N L V K E R S R E P E S S D S D E K V H L A . P S S . S K R K L . E . E . T V E . K K K G . R K D S R I
yHIR2 531 T N V K L T S K V L D Q D V L F Q D .
yHIR1 590 N A G G K G S D G T S N . S I D I . V L S E . E N D F E R M T L N A K L T Q E K . W S . E P T T . C L L Q S D V I P D T

hsCAF1A 868 K T Q E G . V L M E . P E Q Q P L S Q E . V D R K G M F V E I D N R W K . G V E T T Q I K L I K K . Q T A E N E D
ceHILP 651 L P . S . S V Q S P A A . S T E K E . M C . S A P A L A L K L P I P G P Q . R F T L Q V S S D P S . Y I . V E N E T T
mmHIRA 550
yHIR2 550
yHIR1 650 D V . V . E G G S L D D . V L E I R N G E . . . . . R S E Q F S E A L . E D N

hsCAF1A 928 D M M D G E R R . . . . . N E E . . C L M . V I G S P V . I V A N K E N V V . C G D . C I R V R . R . P C G T H V S E
ceHILP 711 V R G . R L S R . K C N R E G . K E W E . L L S S R . T T A E S C D V . C M A C K M S V . S C G R R L R E P
mmHIRA 550
yHIR2 550
yHIR1 686 P T R . E L G Y Q G G K R . . . . . T I E T F I P E V I I C A I . S K D . K C W C P S A N G S . Y E L S Y N Q Q R R P E

hsCAF1A 984 R E D S L P V L G V . E H A A Y A L E N R R G T N M L G K A V T R O P L F D C T E R S D N .
ceHILP 770 P P P S T T H C R G P V A G A T A T L S V W D V R Q V . V A V . E S E S H S L G S D M .
mmHIRA 589 P E V S I S F L E A C G T Y . C L A S I . E L Y C . E Q K K L A P P T N T I Y P L . N P S L R Y S D D I L T R
yHIR2 743 C L G E K . K I K . V T S K Y . V . E R . L F F A . D . L D L K L V R N V P I . P L . N G Q P I H G N K V R I N K
yHIR1 743

hsCAF1A 1037 E I S A D . S E T G . P R V F S A G S I . T E N V S I S C W I Q .
ceHILP 822 V S Q I . . . . . D G H L P . N L S I G K A . C E N P S E T N L V S D .
mmHIRA 649 E N I T . C S . I T K K G . P . V T L S N G D G Y L F D K N M E T W L V S D G W W A Y G S Q Y W D T T N T T G L S S
yHIR2 803 V I R K F R L D G S S C D . L . E V G . P K N V Y K W T R D L G C W S L Y K .
yHIR1 803

hsCAF1A 1070 . . . . . T N V L G R L S E I S D P O I S C S N G T A N S A C P L V R L L K R M R K Q A T P G
ceHILP 858 . . . . . K O P . . . . . Q C . D F E M S L S Q D A M . C S G P L . I I Q . . T S V S G R O F A R E F S .
mmHIRA 708 K A N T D S F N G S E S . E I N E I V . S D I . N D N Q S I I N F L E C K T N D E . N R R K G I K N L O R P A R T E L M K E
yHIR2 708
yHIR1 708

hsCAF1A 1117 V A P Q V K A V K E . Q E E L . H C E . E Q G N P Q E M Q T M L . P V E T L C E G G S E K K R N L N D L S R S
ceHILP 904 V P H V . Q E T L P . Y E D N O . A A . T L S S E . S R W L . P A R Y L V N E G . E Y R L . T C K D L L G P
mmHIRA 768 G F E N . I V N . E B E N K L I . R R . E P E S . S . L . . C I R L S E L G . M D R L N . F Q W L Y D D
yHIR2 768
yHIR1 768

hsCAF1A 1177 G A P M Q V C G L R R S A Y D D V T R . . . . . L K Q A . . . . . A R V A G V A A S T T T K S F .
ceHILP 964 H C S T G S Q W E S . V . G L R R E . V L K E P L E V L G Q N . R F O R F E C Q E . . . . . P R D K
mmHIRA 828 P I S . G G . S . F A D K D F R R N . . . . . K L I A C . D I R Q V O R . T R Y A K . . . . . S .
yHIR2 828
yHIR1 828

```

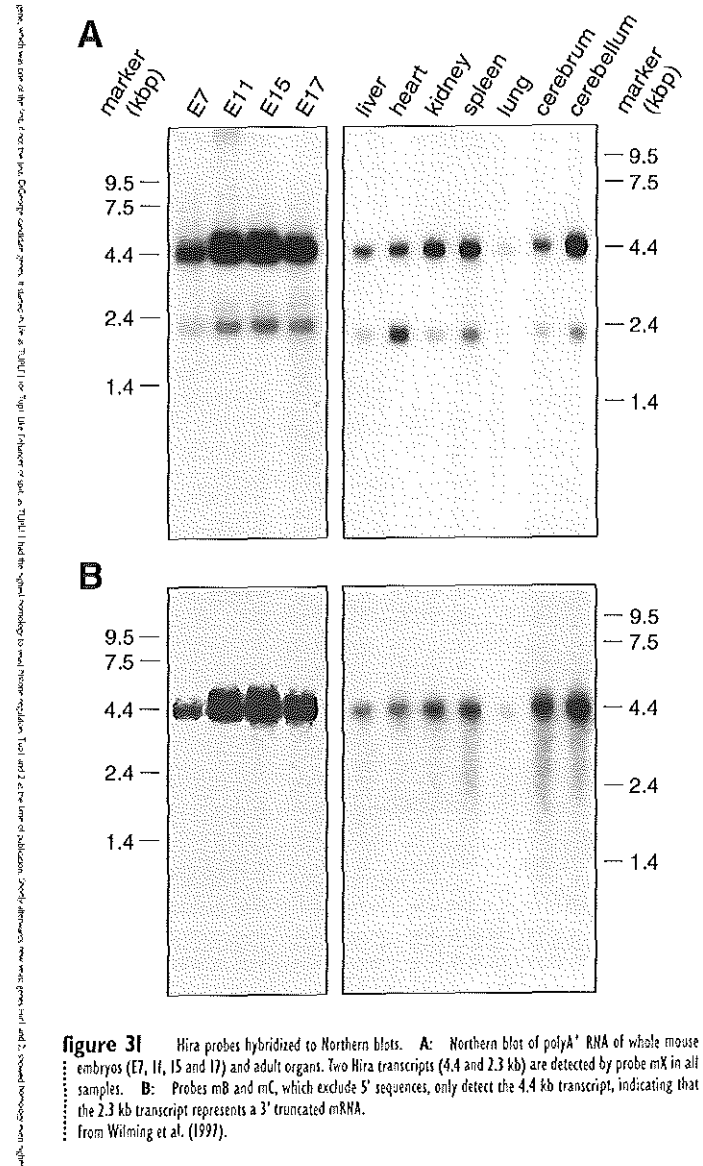
figure 30 Alignment of mouse Hira sequence with hsCAF1A, ceHILP, yeast Hir2p and Hir1p. Identical amino acids are indicated by black boxes, similar residues by gray boxes. The homologies shown are generated by comparing CAF1A, ceHILP, Hir1p and Hir2p to Hira. Three regions can be distinguished with high, low and intermediate homology. Region A (aa 1-371) spanning the WD40 repeats; region B (aa 473-658); region C (aa 744-1015). ceHILP has 291 additional N-terminal residues. From Wilming et al. (1997).

drops (average similarity 18%); C-terminal region C (aa 744 till 1015) shows again an increase in homology (average similarity 35%). In this alignment the C-termini of Hira, ceHILP and Hir2p almost coincide. CAFIA terminates at the end of region B and Hir1p terminates in region C. Mouse/human HIRA alignments and the intron/exon structure of the *ceHILP* gene support the presence of three distinct regions within Hira: mouse/human HIRA comparison showed higher homology in the N-terminal one third (region A) than in the remainder of the protein; ceHILP exon-exon boundaries are spatially remarkably close to those of Hira. The major part of regions B and C are coded for by large single exons in *ceHILP* (Wilson *et al.* 1994; U21322), suggesting that these regions represent single functional domains. Region A contains the WD40 domains, domains that are present in a wide range of different proteins (Neer *et al.* 1994) and are involved in protein-protein interactions (Sondek *et al.* 1996). In the case of yeast Tup1p, after which Hira was originally named, WD40 domains 1 and 2 interact directly with the homeo-domain protein a2p (Komachi *et al.* 1994).

The homology in regions B and C points to conserved domains although their functions have yet to be demonstrated. A clue to their functions might again come from Tup1p. Recently, Edmondson *et al.* reported that Tup1p binds histones H3 and H4 through a domain outside the WD40 region (Edmondson 1996). Strikingly, the CAF complex also binds histones H3 and (acetylated) H4 and is thought to chaperone these core histones to the newly replicated DNA (Kaufman *et al.* 1995). In view of the homology of Hira with CAFIA it might well be that Hira performs an analogous function. Moreover, CAFIA function might be sensitive to gene dosage. CAFIA has been mapped to the Down syndrome critical region at 21q22 (Blouin *et al.* 1996). The presence of three copies of CAFIA in Down syndrome individuals might contribute to the phenotype. CAFIA works in a complex with at least one other subunit (p150). Changes in gene-dosage might perturb the stoichiometry of the complex, hampering its formation. Hira, like CAFIA and Hir1p and Hir2p, is expected to be functional in a complex. Decreased availability of Hira due to haploinsufficiency of its gene may well have an effect on the formation of such a complex.

VIII 3 TWO HIRA TRANSCRIPTS ARE PRESENT FROM GASTRULATION STAGE ONWARD

The presence of *Hira* mRNA in adult mice and embryos was examined by Northern blot analysis of polyA⁺ RNA extracted from embryos at different stages of development and from various adult organs. Probes 12E (cDNA) and mX (figures 28A, B) detect transcripts of 4.4 and 2.3 kb (figure 31A). The 4.4 kb transcript is more abundant in most organs while the 2.3 kb transcript is more abundant in the adult heart. Among the tissues analyzed, cerebellum, kidney and spleen contain the highest levels of *Hira* mRNA, while lung contains lower but detectable levels. We also



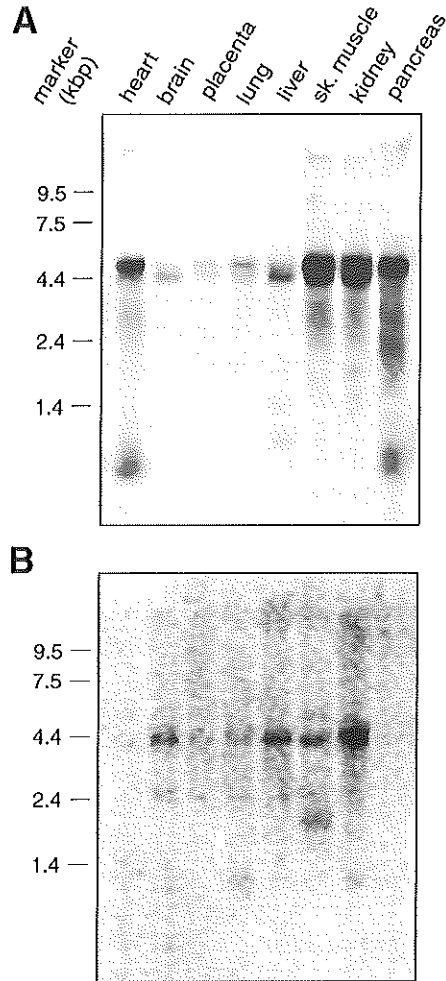


figure 32 Northern blots with Poly(A)⁺ RNA from adult human organs. **A:** Hybridized with probe hE, two transcripts are detected (4.5 and 5.0 kb). Relative abundances vary from organ to organ. The 4.5 kb transcript is most abundant in heart and pancreas. **B:** Mouse probe mC only detects the 4.5 kb transcript and an additional faint band of ~2.4 kb. From Wilming et al. (1997).

FIGURE 32. Northern blots of Poly(A)⁺ RNA from adult human organs. **A:** Hybridized with probe hE, two transcripts are detected (4.5 and 5.0 kb). Relative abundances vary from organ to organ. **B:** Mouse probe mC only detects the 4.5 kb transcript and an additional faint band of ~2.4 kb. From Wilming et al. (1997).

detect *Hira* mRNA in the mouse P19 cell line, the murine neuroblastoma cell line N2A and in the rat pheochromocytoma cell line PC12 (not shown), indicating *Hira* expression in tumors of neuroectodermal origin.

In mouse embryos the 4.4 and 2.3 kb *Hira* transcripts are already present during gastrulation stages (E7) (figure 31A). At E7 the signal is lower than at more advanced stages. At all stages examined the 4.4 kb transcript is most abundant.

Given the length of the human and mouse *HIRA* cDNAs of approximately 4 kb (Halford et al. 1993a, b, c, Scamps et al. 1996), the 4.4 kb *Hira* transcript most likely represents full-length *Hira*. Probes mB and mC (figure 28B) only detect the 4.4 kb transcript on the organ and embryo Northern blots, respectively (figure 31B). This suggests that the 2.3 kb transcript contains sequences from the 5' part of *Hira* only. Thus, the 2.3 kb transcript might represent an MF2-like mRNA (Halford et al. 1993a, b, c).

To compare *Hira* expression in adult mouse organs with expression in adult human organs, we hybridized a human multiple tissue Northern blot with probe hE (figure 28B). Different from what we observed on the Northern blot of adult mouse organs, probe hD detects two transcripts of approximately 4.5 kb and 5 kb (figure 32A). The relative intensities of the two transcripts vary from organ to organ. In heart, lung and pancreas the 5.0 kb transcript is most abundant, while in brain and liver the 4.5 kb transcript prevails. Mouse probe mB on the same Northern blot fails to detect the 5.0 kb transcript (figure 32B). Only the 4.5 kb transcript is detected. However, an additional weak signal at 2.4 kb is seen.

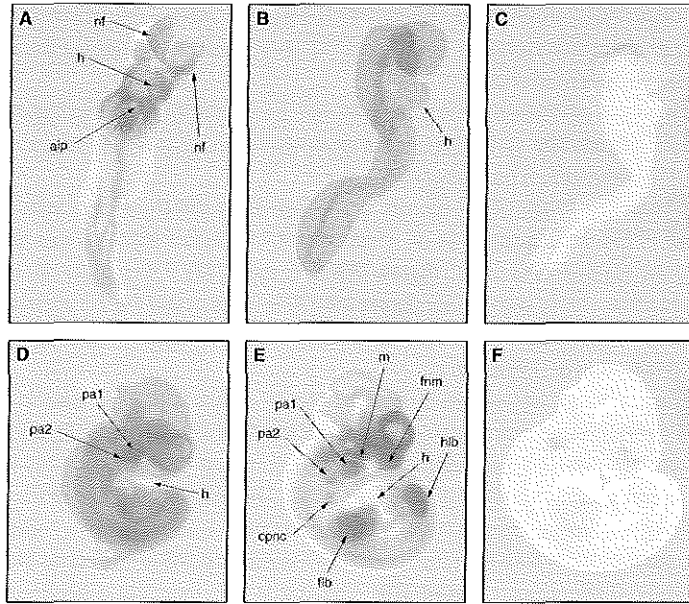
VIII 4 *HIRA* IS UBQUITOUSLY EXPRESSED, WITH ENHANCED EXPRESSION IN PHARYNGEAL ARCHES, NEURAL FOLDS, CIRCUMPHARYNGEAL NEURAL CREST AND LIMB BUDS

The expression of *Hira* in organs (brain) and cell lines (neuroblastoma and pheochromocytoma) of neuroectodermal origin and its expression during organogenesis, raised the possibility that *Hira* plays a role in the development of the brain and the pharyngeal region of the head. Therefore we determined the spatial distribution of *Hira* during mammalian embryogenesis. Whole mount *in situ* hybridization was performed on mouse embryos of 8.5, 9.5 and 10.5 days of gestation, a time-window during which the most critical events in organogenesis are taking place. For these experiments we used single stranded RNA probe mA (figure 28B), that only detects the 4.4 kb transcript.

At E8.5 base levels of *Hira* transcripts are present throughout the embryo, including the heart (figures 33A, B). Higher levels of transcripts are detected in the cranial neural folds (figure 33A). At E9.5 *Hira* transcripts are also widely present (figure 33D). Higher levels of expression are seen in the pharyngeal arches 1 and 2. We do not detect *Hira* transcripts in the heart anymore at this stage.

figure 33 (this is a grayscale thumbnail of the original for quick reference; see the full size color picture on the foldout poster on the back cover)

Whole mount in situ hybridization of mouse embryos with DIG labelled probe mA. **A:** Ventral view of an E8.5 embryo. *Hira* transcripts are present ubiquitously, but higher levels are present in the neural fold. **B:** Lateral view of an E8.5 embryo. The heart is positive for *Hira*. **C:** E8.5 embryo hybridized with a non-complementary (sense) probe as a negative control. **D:** Lateral view of an E9.5 embryo shows enhanced expression in the first two branchial arches. The heart is notably negative. **E:** Lateral view of an E10.5 shows additional expression in limb buds, the circumpharyngeal neural crest, the frontonasal mass and the maxillary component of arch 1. **F:** E10.5 embryo as a negative control. Abbreviations: aip: anterior intestinal portal; cpnc: circumpharyngeal neural crest; flb: forelimb bud; hlb: hindlimb bud; h: heart; m: maxillary component of pharyngeal arch 1; nf: neural folds; pa1: pharyngeal arch 1; pa2: pharyngeal arch 2; Lateral views are shown with the dorsal side of the embryo left and all views show the anterior end up. from Wilming et al. (1997).



Sections of gelatin embedded whole mount embryos reveal *Hira* transcripts in the mesenchyme of the pharyngeal arches (not shown). At E10.5 the expression pattern of *Hira* is similar to that at E9.5 (figure 33E). The highest level of transcripts is in the anterior part (including the maxillary component) of arch 1 and the posterior part of arch 2. In addition, the frontonasal mass and limb buds have a raised level of transcripts. Just posterior to arches 1 and 2 a circumscribed spot with a higher level of *Hira* transcripts is present. This region corresponds well to the region where hindbrain neural crest cells reside before entering the developing gut and migrating to the third and fourth pharyngeal arch. In chicken embryos this region is known as the circumpharyngeal neural crest (Kuratani and Kirby 1991). Control hybridizations with non-complementary (sense) probes are negative (figures 33C, F).

It is of interest that *Hira* transcripts are present in neural folds, where neural crest cells will be generated. In chicken embryos the zinc finger gene *Slug* is also expressed in the cranial folds. Treatment of embryos with *Slug* antisense oligos showed that *Slug* is necessary for neural crest cell migration (Nieto et al. 1994). In mice, the homeobox gene *Msx-3* shows highly restricted expression that

partially overlaps with *Hira* expression (Shimeld et al. 1996). Most of the mesenchymal cells in the pharyngeal arches derive from hindbrain neural crest cells. It thus seems likely that the *Hira* expressing cells in pharyngeal arches 1 and 2 represent hindbrain neural crest cells. Special attention needs to be drawn to the regional differences within the arch mesenchyme. Certain sub-populations of cells either in the anterior part (arch 1) or in the posterior part (arch 2) seem to exist. In a recent study it was shown that *Bmp-7* has an overlapping distribution pattern (Wall and Hogan 1995). It will be interesting to study whether the various gene products (*Slug*, *Msx-3*, *Bmp-7* and *Hira*) function in the same developmental pathway. Another group of *Hira* expressing cells that is evident constitutes of neural crest cells in the circumpharyngeal region. In birds, it has been demonstrated that these neural crest cells are crucial for proper development of the outflow tract of the heart. In mice, these cells express the proto-oncogene *Ret*, and part of this population of hindbrain neural crest cells is on its way to the developing gut or to the superior cervical ganglion (Durbec et al. 1996).

In general, the elevated levels of *Hira* expression overlap with the expression domains of the cellular retinoic acid binding protein gene *Crabp1*. *Crabp1* is also expressed in the frontonasal mass, branchial arches 1 and 2 and the neural folds. Exposure of vertebrate embryos to retinoic acid leads to phenocopy of the CATCH22 phenotype (Vaessen et al. 1990). The overall *Hira* expression pattern seems to be conserved across species (Roberts et al. 1997).

Full-size *Hira* transcripts are present in the developing heart at E8.5, where they are not detectable anymore at E9.5 and E10.5. The probe we used for the in situ experiments does not detect the short transcript, which is predominant in the adult heart. It is possible that between E8.5 and E9.5 the embryonic heart switches from the full-size to the 2.3 kb transcript. The developing heart arises through a complex series of morphogenetic interactions. In the mouse embryo the heart field is induced at E7.5. The straight heart tube is subsequently specified in an antero-posterior sequence (E9.5) to form the various regions and chambers of the looped (E8.5) and mature heart (E15) (Kaufman 1992). The expression in the heart at E8.5

suggests that Hira might play a role in cardiac looping. This is of interest in the context of CATCH22, as the heart is one of the affected organs.

VIII 5 CDNA CLONES 12B AND 12E REPRESENT SPLICED VARIANTS WITH EXONS 3A AND 11A

Two out of the seven *Hira* cDNAs (clone 12B and 12E) contain non-canonical sequences (figure 28A) in positions similar to the human clones HT4 and HT5. To investigate whether these cDNAs represent valid transcripts, we partially characterized the genomic organization of the murine *Hira* gene. The intron-exon boundaries we determined coincide with those in the human as recently published by Lorain *et al.* (1996). Below we will use their numbering of the exons.

Clone 12B contains a 203 bp insert between exons 3 and 4, which encode part of the first and second WD40 domain respectively (sequence in figure 34A). Genomically, the insert is continuous with exon 3 but is separated from exon 4 by an intron. Therefore, clone 12B must have arisen from the use of an alternative splice donor site. The ORF of clone 12B is interrupted by a frameshift, both in cDNA and genomic clones. The next potential translation initiation site is

A

```

1  gtattacaga  gtggcccttt  gcaggctttg  cgtgttgcca  gagtgcceaa  50
51  ggttageaca  tgtgtttgta  attagaaga  taaggcccg  tggcgctta  100
101 actcccctcg  gcctcgtgct  tgaaggca  gatgttgct  gcacacctgg  150
151 gtgagttatt  cggcagaaga  ctggcgccc  acgaaccatt  tggacgtctg  200
201 cga  203

```

B

```

1  gacttttgcc  aaccgttaat  ggggtattat  atccatttgt  ctgaatggag  50
51  ctctgtggaa  gccacagaca  ggtcagcagc  tateccactta  gtgatgctgt  100
101 tgctggg  107

```

figure 34 Nucleotide sequences of the extra *Hira* exons. **A:** exon 3A from clone 12B. **B:** exon 11A from clone 12E. Stopcodons in the *Hira* reading frame are underlined. Accession numbers: EMBL X99713 and X99722, respectively. From Wilming *et al.* (1997).

the methionine codon at nucleotide position 510, with a Kozak score of 0.67, or the methionine codon at nucleotide position 546 with a Kozak score of 0.78.

The 3' terminus of clone 12E has an insert of 107 nucleotides between exons 11 and 12 (sequence in figure 34B). Exon 11 encodes the last WD40 domain. The 12E insert causes a frameshift in the ORF, with stop codons as a result. PCR amplifications with primer sets p5-p8 and p6-p8 (figure 28B) on mouse genomic DNA, show that the 12E insert is located within an intron. With primers p5 and p8 we can amplify a 1.3 kb fragment, while with primers p6 and p8 we can amplify a 0.6 kb fragment. The insert is therefore separated from exon 11 and 12 by 0.5 kb and 0.33 kb introns, respectively.

Genomic sequences show that the 25 bp MF2 insert is positioned between exons 2 and 3. The extra sequences are designated by us as exon 2A (MF2), 3A (12B) and 11A (12E). BlastN and BlastX database searches with exons 3A and 11A do not yield any significant homologies.

Through RNase protection assays, we examined the presence of cellular mRNA species containing exon 3A. We detect full length protected fragments with probe mX (figures 28B, 35B) in all RNA samples analyzed (figure 35A). This indicates the presence of canonical *Hira* transcripts. In addition, two smaller protected fragments are detected with this probe showing the existence of mRNA species that contain exon 3A. Full length protection was also found with probe mY (figure 28B) that contains exon 3A sequence (not shown). This also indicates the presence of transcripts containing exon 3A. None of the negative controls with sense probe or tRNA target yield any protected fragment.

To show the presence of mRNAs containing exon 11A we performed RT-PCR experiments with various primer sets on mouse embryo and adult organ RNAs. We used primer sets p5-p7, p5-p8, p6-p7 and p6-p8 (figure 28B) on poly-dT and random primed cDNA from total or polyA⁺ RNA samples. With all four primer sets we can amplify fragments containing exon 11A sequences, both in adult organs and embryos. The presence of transcripts containing exon 3A was likewise confirmed by RT-PCR with primer sets p1-p3, p1-p4, p2-p3 and p2-p4 (figure 28B) on the same samples. Exon 3A is present in all samples (not shown).

Recapitulating, we find that two *Hira* mRNA species (2.3 and 4.4 kb) are present in embryos and adult organs. Both embryonically and at the adult stage (averaged over the individual organs) the 4.4 kb transcript is expressed at higher levels than the 2.3 kb transcript, without dramatic variations in the ratio between the two. Between adult organs the ratio of the expression of the 2.3 kb and 4.4 kb transcripts varies. The smaller 2.3 kb transcript is the predominant mRNA in the adult heart. Northern blot analysis reveals that this transcript lacks the sequences of the 3' end of canonical *Hira*. This mRNA thus encodes a protein that consists of mainly WD40 domains. The 2.3 kb transcript could represent the transcript coding for the truncated 12E or MF2 coded polypeptides. The 4.4 kb transcript must represent the full length *Hira* transcript. Scamps *et al.* mention only one full length

transcript in fetal mouse RNA (Scamps *et al.* 1996). Most likely, they used 3' probes.

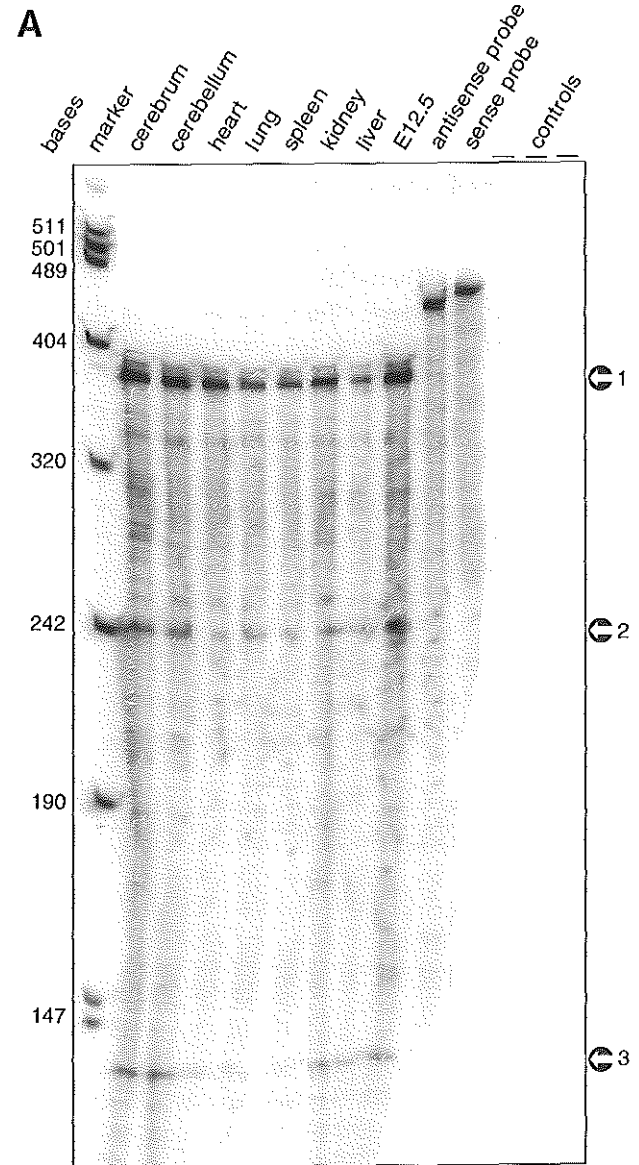
Until now, two *HIRA* transcripts have only been reported for human fetal liver (Halford *et al.* 1993a, b, c, Lamour *et al.* 1995). Though these transcripts were labelled 3.4 and 3.2 kb (Halford *et al.* 1993), and 4.5 and 4.2 kb (Lamour *et al.* 1995) we assume that these are identical to our 5.0 and 4.5 kb adult transcripts. In the adult heart and pancreas the 5.0 kb transcript is predominant, whereas the 4.5 kb transcript is predominant in brain. Like the 2.3 kb transcript in the mouse, the 5.0 kb variant is not detected by a *Hira* probe that excludes WD40 repeat coding regions. One explanation could be that the 5.0 kb transcript represents an mRNA that encodes a *HIRA* like protein with C-terminal WD40 domains. This protein would have an N-terminal extension, which could be similar to the N-terminal extension of ceHILP. This part of ceHILP, which has no equivalent in *HIRA*, CAF1A, Hir1p or Hir2p, bears very high homology to *C. elegans* proteins that are similar to *S. pombe* SDS22p and *L. monocytogenes* Internalin (U39849, P45969). Yeast SDS22p positively modulates protein phosphatase 1 (Ohkura and Yanagida 1991) and Internalin is required for the internalization of *L. monocytogenes* by eukaryotic cells (Mengaud *et al.* 1996). Alternatively, the 5.0 kb transcript encodes a protein with N-terminal WD40 repeats followed by as of yet unknown sequences not present in canonical *HIRA*.

The insertion of exon 3A into the *Hira* mRNA disrupts the canonical *Hira* open reading frame. An N-terminally truncated protein starting from either of the next two in-frame translation start sites lacks the first WD40 domain. This is comparable with the protein coded for by the MF2 *Triple 1* (*Hira*) cDNA, which lacks half of the first WD40 domain due to a 25 bp insert in the mRNA. In *Triple 1* the array of WD40 domains appears to influence the binding specificity of the individual domains: in the context of other WD40 domains, a single domain has different binding specificities than in isolation (Komachi *et al.* 1994). The lack of a WD40 domain is therefore very likely to alter the affinity of *Hira* for other proteins. The insertion of exon 11A introduces stop codons, yielding a polypeptide consisting only of WD40 domains. This truncated polypeptide could compete with full length *Hira* for binding to other proteins. This putative truncated form of *Hira* resembles CAF1A which also consists mainly of WD40 repeats.

In conclusion, at least four different splice forms of *HIRA* can be identified in humans and mice. This complex expression pattern of the mammalian *HIRA* gene is regulated in an organ- and developmental specific manner and arises from alternative splicing and exon skipping.

H*IRA* is being analyzed for point mutations and microdeletions by several groups that are looking for a cause for the CATCH22 phenotype in apparently non-deleted CATCH22 patients. The new exons 2A, 3A and 11A should be included in these screens. Further analyses of the human 5.0 kb transcript should help in identifying more *HIRA* sequences to check for mutations.

S10287 from The European Bioinformatics Institute (EMBL) database, accession number U00001.1. The sequence of the HIRA gene is available in GenBank, accession number U00001.1. The sequence of the HIRA gene is available in GenBank, accession number U00001.1. The sequence of the HIRA gene is available in GenBank, accession number U00001.1.



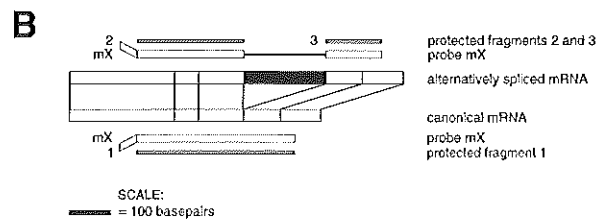


figure 35 Result of RNase protection assay and diagram of the probes used. **A:** Protected RNA fragments resulting from an assay with probe mX and the target RNAs shown. Controls were RNAs hybridized to non-complementary (sense) RNA probes or tRNA as a target for antisense probes. Arrowheads point to probe mX (sense and antisense) and the protected fragments: (1) full length protected fragment representing the canonical transcript; (2) and (3) indicate the presence of transcripts with exon 3A. **B:** Schematic drawing of probe mX and the relevant fragments of two possible target transcripts. Gray boxes indicate canonical exons, the black box indicates exon 3A. The upper and lower lines indicate the expected size of protected fragments and their corresponding number in **A**.
from Wilming et al. (1997).

refer and later. There is enhanced recognition in the 5' end and the 3' end and amino phosphorylation. Together with the chromosomal location and its possible location (gene or chromosome region), we conclude that H19 is an excellent C192 candidate gene.

ix

Discussion & Conclusions

Commonly deleted in the majority of patients with one of the chromosome 22 deletion syndromes is a region considerably larger than the SRO. Most CATCH22 deletions extend over a distance of more than 3 Mb, usually including the *sc11.1B* locus (Desmaze *et al.* 1993a, b, Halford *et al.* 1993c, Lindsay *et al.* 1995, Morrow *et al.* 1995, Carlson *et al.* 1997a, b), as was also shown in our patient series. According to Halford *et al.* (1993c), a deletion not including the *sc11.1B* locus was found in 7 out of 91 DGS and 6 out of 16 VCFS cases. In six of the seven DGS cases, the deletion includes the D22S134 (*sc4.1*) locus, whereas in the six VCFS cases, the deletion does not extend distally to this locus. The only DGS patient who probably had a deletion not extending so far distally is the child of the father who's cell line is encoded GM05878. In VCFS the most centromeric distal boundary is found in GM00980. This means that haploinsufficiency for a gene or genes in the region between the GM00980 and GM05878 breakpoints, and perhaps also in the adjacent distal region, might play a specific role in the pathology of DGS. Furthermore, genes located more distally but within the commonly deleted region, such as *COMT* (Dunham *et al.* 1992, Desmaze *et al.* 1993a), *T10* (Halford *et al.* 1993b) and the zinc finger gene *ZNF74* (Aubry *et al.* 1993), might contribute to the clinical phenotype in CATCH22 patients. Occasionally 22DS patients with distal deletions are found, including genes like *TBX1* and *COMT* (McQuade *et al.* 1998). Although *ZNF74* was found to be deleted in 23 out of 24 DGS patients (Aubry *et al.* 1993), our FISH analysis now shows that the gene maps close, and possibly even distal, to *sc11.1B*, i.e. far outside the SRO. This is not to say that in the more distal regions there could not be contributing genes. For example, Der(22) syndrome patients suffer from severe developmental delay and craniofacial and cardiac abnormalities after meiotic malsegregation of the constitutional t(11;22)(q11;q23) translocation. The genotype results in duplication of the proximal 22q11 region and mapping of the breakpoint shows it to map in the same region as the distal breakpoint of 1.5 Mb VCFS deletions, i.e. between *RANBP1* and

ZNF74 (Funke *et al.* 1998). The overlap in phenotype suggests that anywhere in this region one or more dosage dependent genes can be expected.

Family members with apparently identical 22q11 deletions can be either severely, mildly or not at all affected. In literature several such cases have been described (Holder *et al.* 1993, Devriendt *et al.* 1997). The ADU/VDU patients, which have a balanced translocation in stead of a deletion, are also an example of phenotypic heterogeneity. Cloning of the breakpoints in both mother and daughter has shown them to be identical (Budarf *et al.* 1995). Yet, the mother is asymptomatic (although in hindsight she may have weak facial dysmorphism) and the child has partial DGS. In two of the DGS cases studied here, the deletion was inherited from a parent who does not have the DGS or VCFS phenotype. But the mother of one patient has had an operation for tetralogy of Fallot, and the father of the other patient suffers from psychiatric illness but has no physical complaints. Inheritance of a 22q11 deletion with increased severity of the clinical phenotype in the second generation has been reported before by others (Wilson *et al.* 1991, 1992b, Desmaze *et al.* 1993a, Driscoll *et al.* 1993). Our findings once more indicate that if an individual with isolated conotruncal heart defects is found to have a 22q11 deletion, there is a risk of more complicated malformation in the offspring. These malformations can spread across organs or tissues, i.e. when the parent had a mild conotruncal defect, the proband doesn't necessarily have a more severe form of that specific defect. It can have other or additional affected organs. Here, more complicated means the total picture is more severe with respect to quality, quantity and/or variation. Embryologically the source of these malformations is the same, namely neural crest related. Differential effects of apparently identical 22q11 genotypes could be the result of environmental effects during development, subtle genetic differences (even between identical twins), stochastics, or any combination of the above.

A small number of DGS and a large number of VCFS patients have been described in whom no hemizygoty was found for any of the loci investigated (Halford *et al.* 1993c). The prevalence of such cases may depend on the strictness of diagnostic criteria (Driscoll *et al.* 1993). In our series, one DGS case, four VCFS cases and the cell line GM03479 revealed no hemizygoty. Patient DG-Ro9 (table 5) has the major features of DGS but he has a dysmorphic mentally retarded brother with no DGS features, indicating that the malformations might have another basis in this family. All the VCFS patients included in table 5 have the major features of the syndrome. A genetic etiology is suggested by the fact that, in all four cases, family members have VCFS dysmorphic features. An explanation could be that either there is another VCFS locus or the mutation in 22q11 cannot be detected with our present methods. An indication for another DGS locus comes from patients with a deletion in 10p13 (Greenberg *et al.* 1988a, b, Monaco *et al.* 1991) (see also later). Alternatively, the disease does not have a genetic basis in non-deleted cases.

table 5 Clinical features of four VCFS patients and one DGS patient without detectable 22q11 deletion. +: presence of clinical trait; -: absence of clinical trait; nd: no data available; (a): mother and daughter; (b): sister has VCFS dysmorphic features but no cardiac defect; (c): mother has prominent nose and cardiac murmur; (d): sib has mental retardation, dysmorphic face and deafness.

	VCF-Ro2	VCF-Ro3	VCF-Ro5	VCF-Ro6	DG-Ro9
Cardiac defects	+	+	+	+	+
Typical facies	+	+	+	+	+
Palatal abnormalities	nd	nd	+	+	+
Learning disabilities	nd	nd	+	nd	nd
Clinodactyly	nd	+	+	+	-
Decreased Lymphoid tissue	nd	nd	nd	nd	+
Other	(a)	(a)	(b)	(c)	(d)

Genetic studies of DiGeorge syndrome have been based on the assumption that the critical gene is located at a balanced translocation breakpoint or in an SRDO. The description of new balanced translocation breakpoints and interstitial deletions associated with CATCH22 phenotypes necessitate the revision of this assumption. The definition of a DiGeorge syndrome critical region by the various groups is confusing and conflicting. The Philadelphia MDGCR comprises contig 1, 2 and part of 3 (distal boundary determined by a t(15;22) translocation). This group recently reported a balanced t(21;22) translocation, disrupting the *CTLCL1* gene, in a patient with only facial VCFS features. The suggestion was made that the truncation of *CTLCL1* leads to the features in this patient and that mutations/deletions of other genes must be responsible for cardiac, thymic and parathyroid features of the CATCH22 phenotype (Holmes *et al.* 1997). The boundaries of the European SRDO (represented by our contig 3) are defined by the deletion breakpoint in patient G (Levy *et al.* 1995) and the unbalanced t(11;22) translocation breakpoint in GM00980. The centromeric deletion breakpoint in patient G has not been cloned yet, but I map it between *DGSI* and *SLC20A3*. To add to the confusion, an interstitial deletion distal in the commonly deleted region has been reported in a Japanese patient with a (partial?) Takao syndrome phenotype. This could point to the existence of a second DiGeorge syndrome critical region (Kurahashi *et al.* 1996). This second DGCR is located telomeric of contig 5. It will be important to identify genes in this SRDO2. Furthermore, position effects cannot be excluded. In campomelic dysplasia and aniridia, *SOX9* and *PAX6*, respectively, can be influenced by translocation breakpoints up to 250 kb distant from the disease causing gene (Wirth *et al.* 1996, Fantes *et al.* 1995). A similar effect could play a role in CATCH22. It is also possible that as a result of the CATCH22 deletion a heterochromatic region (e.g. close to the centromere) can exert its influence onto former distal regions, thereby silencing

the genes in these regions. In this model it is even possible that the genes of major effect are located outside the commonly deleted region. Support for this theory may be found in the description of four monozygotic twin pairs with interstitial 22q11 deletions (Miller *et al.* 1983, Goodship *et al.* 1995, Fryer 1996, Yamagishi *et al.* 1998). In these cases there is important variation in the phenotype within single twin pairs. Since genetically the individuals within each twin pair are virtually identical, the explanation for this discordance could be found in position effect variegation. The extent of this proposed variegation could vary between the individuals of a twin pair (due to epigenetic effects), thereby silencing genes over varying distances along the chromosome. Similar position effects could play a role in the above mentioned unbalanced translocations t(15;22) and t(21;22). In these translocations 22q11-qter moves to the p arm of an acrocentric chromosome, regions known to be highly heterochromatic. It is very well possible that the spread of heterochromatin from the p arm of the host chromosome to 22q11 is involved in the silencing of genes close to the translocation breakpoint. Since there appear to be at least two SRDOs, one cannot exclude that both position effects and gene deletions can lead to the CATCH22 phenotype. Considering the fact that there are many phenocopies of the CATCH22 phenotype, deletion or silencing of different genes on 22q11 could lead to similar phenotypes.

There may be a strong correlation between the CATCH22 and 22q11 deletions, but there are reports of correlation with other chromosomes, most notably 10p (Greenberg *et al.* 1988a, b, Monaco *et al.* 1991, Obregon *et al.* 1992, Shapira *et al.* 1994, Devriendt *et al.* 1995, Schuffenhauer *et al.* 1995, Daw *et al.* 1996, Lipson *et al.* 1996, Pignata *et al.* 1996, Hsu *et al.* 1997). Indications are that there are two different loci on 10p, since there is not one single SRO but two close together (Dasouki *et al.* 1997, Gottlieb *et al.* 1998). I think these loci should not be neglected; they may hold important clues as to which genes from 22q11 are involved. Comparing genes in these alternative loci with those in or around 22q11 may reveal genes whose products are involved in the same (developmental) pathway, or maybe even homologous genes.

This would then strengthen the position of the 22q11 gene as a candidate gene and enable the researcher to focus on the most likely causative gene.

From above considerations we tend to conclude that from the position of a gene within the deletion region one cannot deduce whether a gene is crucial or not. All the genes described in the Results & Discussion chapter of this thesis map within the commonly deleted region and could thus contribute to the phenotype. Is it a prerequisite that the gene product is sensitive to gene dosage? I can think of two alternative mechanism that would fit the genotype-phenotype relationship we see in CATCH22, so the answer must be no. One theoretical possibility is that the gene on the other chromosome has to be rendered nonfunctional as well, by for example a point mutation, to get the phenotype. These mutations may have escaped our attention because we simply do not know where to look or we may not have performed mutational analysis on the gene in question yet. As of yet, inheritance patterns do not indicate that this double hit mechanism is a plausible explanation. A second possible, and more plausible, mechanism is that the gene in question does not get transcribed unless it pairs with its partner on the homologous chromosome. It has been reported in literature that some genes need to pair with their homolog if transcription is to take place. In either case the gene product in itself is not sensitive to dosage, it is simply absent by lack of transcription. Once we have pinpointed the 22DS gene(s) transcription and protein studies will clarify the true mechanism.

The pleiotropy of the DiGeorge genotype and phenotype makes it unlikely that only one gene is etiologic for the syndrome: it is most likely that the phenotype is the result of haploinsufficiency of several genes. To be fair, one of the arguments supporting this idea, the fact that deletions associated with CATCH22 are invariably large, is not necessarily incompatible with a one-gene hypothesis. It is possible that the underlying deletion mechanism dictates that deletions terminate in discrete, widely separated regions (low copy repeats, "weak" regions ?) in stead of at random positions. In that case the large size of the deletions would have no relation to the phenotype but be just

a side effect of the mutational mechanism. There is some evidence that there are regions in 22q11 prone to breaking. Around the DGCR there are regions of paralogous groups of genes and sequences that, owing to the high homology between them, could give rise to recombination events resulting in interstitial deletions (or translocations when matching paralogs on another chromosome) (Halford *et al.* 1993a, Eichler *et al.* 1998). In this light it is interesting to note that two deletion syndromes I described in the introduction, namely Angelmann Syndrome and Smith Magenis Syndrome, have similar regions surrounding their chromosomal regions (Eichler *et al.* 1998). Apparently it is a feature of pericentric regions. The Cat Eye Syndrome region is probably the victim of the same type of chromosomal weak spots as 22DS (Eichler *et al.* 1998). Some CES duplications (Type II CES) overlap with the CATCH22 region and their breakpoints co-localize with DGCR breakpoints (McTaggart *et al.* 1998). Indeed, Edelman *et al.* (1999) recently described a common molecular basis for chromosome 22q11 rearrangement disorders (CES, 22DS, der(22) syndrome). Low copy repeat regions named LCR22s (containing (pseudo)genes like *GGT* and *BCRL* and repeats) apparently mediate translocations, deletions and duplications in all three syndromes. These types of mechanisms act on a large scale, but the fact that the DGCR region is so difficult to clone suggests that on a smaller, more local scale there are stability issues as well. As described we, and others, have had difficulties with cosmid walks either due to lack of overlapping clones or matches to scrambled clones. Genes located in these regions are *UFD1L*, *CDC45L* and *TRAIN-MODEL* (M51 contig) and *HUDDL*, *GPIBB* and *TBX1* (M56 contig). Could this point to involvement of mutations (microdeletions, point mutations) in one or more of these genes in CATCH22 etiology?

Over the course of many years the CATCH22 deletion region has been the subjected to intense scrutiny. This resulted in the isolation of many genes and the production of more than 1.2 Mb of sequence. Ongoing efforts will undoubtedly result in further isolation of genes and detailed expression studies of new and known genes. Despite all the effort, to date not one gene located in the critical region stands out as the obvious candidate gene (but: see **chapter X-2-5**); however, we can make some assumptions on which genes are, on the surface, the likely candidates (see below). The major problem is that not only do we do not know whether one or more genes are involved, there is also the possibility that changes in the chromosomal structure may change regulation of genes up- or downstream of the deletion.

IX 2 OUTLOOK

IX 2 1 BIOINFORMATICS AND GENOME PROJECTS WILL BE
A MAJOR BOON TO MEDICAL GENETIC RESEARCH

If we want to study the suggestion that genes outside the DGCR are causative for CATCH22 or contribute to the phenotype, the road ahead may be shorter than one would initially think. Telomeric of the DiGeorge region many genes have been described, but mostly from heavily studied clusters like Immunoglobulin Lambda Light Chain and BCR. Reports on genes in the intervening regions of 22q are few and far in-between since these regions have not been targeted for systematic study. Fortunately there are several future developments that will aid research in the field of medical genetics and functional genomics. Singularly the most important is the completion of human genome sequencing, which is projected to be early next century (rough shotgun sequence finished spring 2000, complete 99.99 % accurate sequence around 2002). But since chromosome 22 was one of the first ones to get sequenced (because of its small size and the number of genetic diseases associated with it), the chromosome relevant to this thesis is actually the first complete human chromosome to be finished (it was completed in 1999; *Dunham et al. 1999*). As our studies showed, having the genomic sequence available is extremely helpful in locating new (potential) genes (which, as should be clear by now, does not mean that we can incontrovertibly identify the disease gene). Sequencing efforts are not only concentrating on the human genome. Mouse genomic sequencing projects are progressing and will yield invaluable information on the location of previously undetected genes and gene regulatory regions by comparing syntenic regions of mouse and human. For mouse chromosome 16, which contains the syntenic region of human chromosome 22q11, such a sequencing program is under way, as described by *Lund et al. (1999)*. Gene prediction algorithms are constantly improving (the more experimental data becomes available the larger and better the training sets) so that we get better in making reliable predictions where we have no homologies to existing genes. It is questionable however, whether gene prediction will develop fast enough to keep abreast of experimental discovery of genes. Mass gene identification schemes like EST and gene signature (*Okubo et al. 1992*) sequencing already yield very valuable information on the existence of genes where no other significant homology is detectable, often supported by gene prediction programs. Ongoing efforts in this field will increase the value of ESTs in gene identification. An interesting development that will certainly prove to be a tremendously powerful tool for linking genes with disease is the production of expression libraries. One of the first complete libraries available is one for the yeast *S. cerevisiae*. For complete libraries it is essential to have the complete genomic sequence, as we have for *S. cerevisiae*. An expression library consists of arrays of unique DNA or PNA¹⁾ fragments, each unique fragment representing a single ORF, transcript or gene. The complete yeast library represents the total of ORFs, both

predicted and confirmed, on the yeast genome. It is in determining which fragments to synthesize that the yeast genome sequencing project comes in. Libraries can also be made from data obtained from EST projects or other mRNA based projects. This yields sets that are specific for, for example, tissues, medical conditions or developmental stages. Gridded arrays of these DNA/PNA¹⁾ fragments on filter, sorted by chromosome, tissue, developmental stage or any other criterion, can be hybridized with total labelled mRNA from wild type and mutated or experimentally treated cells. Subtracting the hybridization result of the wild type from that from the cells under investigation will reveal changes in the expression of each gene in the set. If polysomes are used as the mRNA source, one can even read translational information from these experiments. Already the oligo arrays are available as so-called micro-chip arrays (on coverslip or microscope slide size glass substrates), which offers obvious advantages in economy of use. Moreover, they are easier to use in automated analysis (scanning with a (modified) fluorescence microscope). Only partial human expression libraries, produced from UniGene EST data, are currently available commercially, but complete libraries like those for yeast will certainly become available in the first decade of the next century.

Clearly the future of genetic and genomic research is bright, in no small part thanks to what is arguably the most important endeavor ever undertaken in biological sciences: the Human Genome Project. It will lay the foundation for decades of research and, together with other genome projects (most notably those for pathogens), provide a major impetus to molecular medicine. It also speeds up the first and probably most time-consuming stage of many medical genetic research projects: finding and isolating the gene(s) (or even chromosome) relevant to the particular disease one is working on. At present, finding a gene responsible for a disease still starts with

¹⁾ PNA = Peptide Nucleic Acids. These have an amide backbone instead of a sugar backbone. As a result the backbone is electrically neutral, lacking the negative phosphates. This in turn makes for less repulsion between molecules to be hybridized, more stable PNA:DNA or PNA:RNA duplexes and milder (low-salt or salt-free) hybridization buffers. The molecules are also more stable: resistant to many stripping-hybridization cycles when fixed to filters or glass and refractory to nuclease activity.

linkage analysis (statistically linking probes or markers to the disease, an arduous task that can take years if there is no deletion to work with) and, once successful in that field, constructing genomic contigs of the relevant area. These then have to be sequenced or subjected to a variety of other techniques to identify genes, after which the candidate genes are subjected to mutational analysis. Arriving at this stage takes years; years that can be saved in the near future, as the sequence and most of the genes in it will be readily available and identifiable. In combination with gridded expression libraries it will make identifying candidate genes both easier and faster. If there is still a need for linkage analysis, the availability of the total genomic sequence yields a sheer infinite number of probes which speeds up the process considerably. Techniques like High Throughput Genotyping make screening for mutations (once genes have been identified) much easier. Of course, having the sequence of genes handed to the researcher on a silver platter, just shifts the emphasis from isolating the gene to determining what the gene product actually does and if, why and how it is responsible for the phenotype.

In the field of genomic analysis, automatic analysis and annotation is a project in constant development. The ACeDB package and associated analysis systems set up at the Sanger Centre in Hinxton (UK) and the Genome Sequencing Center in St Louis, Missouri (USA) are coming a long way in this, but final human intervention is still, and probably always will be, necessary. Systems like these not only allow a great deal of automation, they also offer opportunities for integration of data. Thanks to the take-off of the Web and related technologies, we are now able to link to various sources, offering data on for example mapping and expression from the annotated sequence map (Helt *et al.* 1998 (and Genome Research 8 (1998) in general), Bard *et al.* 1998).

IX 2. 2 REVERSE GENETICS: BE AWARE OF THE PITFALLS OF USING THE MOUSE AS A MODEL ORGANISM

Currently the most popular technique for trying to identify the function and phenotype of a given gene is the construction of mutated organisms, mostly mice but also yeast, *Xenopus* and Zebrafish. We

too think knock-out studies in mice may provide some clues as to which genes are involved in the CATCH22 phenotype. We are currently analyzing the effect on development of a null mutation of the murine *Hira* gene (chapter IX-2-4). If it turns out that more than one gene is involved, crosses between various knock-out lines are necessary. Do we want to test if chromosomal structure effects play a role, then large deletions (in excess of several hundred kb) are necessary. Detailed deletion experiments on multi- and single gene scale are currently under way (Baldini *et al.* 1998). A complicating factor in this matter is that the genomic organization of the mouse syntenic region on chromosome 16 is different on the macro level (Botta *et al.* 1997, Puech *et al.* 1997). Initially it looked like the gene-order was conserved between mouse and human (Bucan *et al.* 1993) and on a certain level it indeed is: between *HIRA* and *RANBP1* the order of the 10 genes tested is identical between the two organisms. But this block of genes is rotated with respect to the *IDD* to *GSCL* block of genes, so that in mouse the *Ranbp1* gene, in stead of *Hira*, is close to *Gscl*. The *IDD*-*GSCL* block by itself is again conserved, but surrounding genes (*DGCR6*, *HCF2*) are displaced up to several megabases (Puech *et al.* 1997). In the results and discussion section I already mentioned that the *CLTCL1* gene is absent from mouse chromosome 16, while the human equivalent of mouse *Tsk1* has degenerated into a pseudogene. Thus, trying to mimic large human CATCH22 deletions in the mouse requires a cautious approach. Within certain limits (the *HIRA*-*RANBP1* block for example) it should be feasible to construct faithful copies of human deletions, deleting exactly the same genes as in human. But deletions spanning the equivalent of the common deletions in human patients may be too dissimilar, deleting genes in mouse that are not deleted in human and *vice versa*. Very recently Lindsay *et al.* constructed mice strains with a deletion of 1.2 Mb of the equivalent of the DiGeorge region, using the cre-loxP system (Lindsay *et al.* 1999). They were able to reproduce CATCH22 like heart defects (related to maldevelopment of the fourth arch arteries) in hemizygous mice and could rescue deleted mice by crossing with mice carrying a duplication of the deleted sequence. The deletion effectively removes the sequence for genes *Dgsi*, *Gscl*, *Ctp*, *Drcg6*, *Ranbp1*, *Arvcf*, *Comt*, *T10*, *Tbx1*, *Gp1bb*, *Cdcrell1*, *Tmvcf1*, *Cdc45l* and *Ufd1l*. From their results it is clear that one or more of the genes in the deletion is dosage dependent and involved in neural crest development in the fourth arch. But it is also clear that other genes, outside this deletion, are responsible for the other phenotypes related to parathyroid, thymus and the craniofacial area. No abnormalities with respect to these structures was observed in the deleted mice.

The fact that the 22q11 region is reshuffled to such an extent, and that the region is distributed over three chromosomes in mouse (6, 10 and 16) suggests that already during evolution the region was unstable and prone to rearrangements. This fits well with the observation that human 22q11 is very unstable. The evolutionary breakpoints may well be related to clusters of breakpoints ("weak spots") on 22q11. It would be interesting to see what the situation is on mouse 16 or in the areas of the evolutionary breakpoints on mouse 6, 10 and 16. Currently the only evolutionary breakpoint

determined at the nucleotide level is one of the breakpoint between the Siamang gibbon (*Hylobates syndactylus*) syntenic region of human chromosomes 22 and 16. This breakpoint is located in human 22q13.1 (Cordelia Langford, PhD thesis (in preparation)). Most primates have the human chromosome 22 syntenic region on one chromosome (in one or two contiguous blocks). Great apes (Gorilla *Gorilla gorilla*, orangutan *Pongo pygmaeus*, chimpanzee *Pan troglodytes*) have a single "human chromosome 22", in their case chromosome 23, while lesser apes (lemur *Eulemur fulvus mayottensis*) distribute human 22 in two blocks over two chromosomes (Jauch et al. 1992, Wienberg et al. 1992, Koehler et al. 1995a, b, Consigliere et al. 1996, Hameister et al. 1997, Richard et al. 1996, Bigoni et al. 1997, Morescalchi et al. 1997, Muller et al. 1997). Cattle (*Bos taurus*), horse (*Equus caballus*), deer (Indian muntjak *Muntiacus muntjak vaginalis*), pig (*Sus scrofa domestica*) and carnivores (American mink *Mustela vison*, harbor seal *Phoca vitulina*, dolphin *Tursiops truncatus*, dog *Canis familiaris*, cat *Felis catus*) give us a glimpse of what was probably an early arrangement of chromosome 22: almost invariably a chromosome 22 block is adjacent to a chromosome 12 block (Hayes 1995, Johansson et al. 1995, Rettenberger et al. 1995a, b, Yang et al. 1995, Chowdhary et al. 1996, Goureau et al. 1996, Fronicke et al. 1996, 1997a, b, Raudsepp et al. 1996, Hameister et al. 1997, O'Brien et al. 1997b, Wienberg et al. 1997, Bielec et al. 1998). In an evolutionary distant past, chromosome 22 and 12 were probably one.

The mouse knock-out system is undeniably a powerful one, but there are some caveats. An important one is that due to redundancy the deletion of one or both copies of the gene of interest could have no measurable effect. Though actually very important and helpful, most of the time reports on these "negative" results are not published, hence giving the impression that knock-outs are more successful than they really are. Another problem regarding phenotype is that a mouse which may be regarded as having no phenotype, actually has one, but we don't know how to recognize it. There are only so many traits one looks at, and the proficiency in detecting anomalies depends heavily on the context of the research and one's particular area of expertise. This problem can partly be alleviated by enlisting the help of independent mouse experts. A further caveat is that the phenotype may escape our attention because the laboratory environment where the mice live will not bring out that particular trait (certain social skills, searching for food, etc.). Short of releasing and monitoring the mice in a more natural biotope, which is logistically and financially next to impossible, we cannot do much about this. Lastly, and probably most importantly, a phenotype can be very dependent on the genetic background, as graphically illustrated by the widely variable phenotypes of *EGFR* knock-out in three different strains of mice (Threadgill et al. 1995). In this respect it is very important to perform the studies described above, on the collinearity of 22q11 genes between mouse and human, on other strains or crosses of mice as well. I would be surprised if the results were identical to those for the published crosses.

A possibly very fruitful way of finding crest genes is turning to the zebrafish (*Danio rerio*). The developmental biologist's pet organism of the late 90s and beyond presents with a list of mutations (natural or induced) that can be traced back to defects in crest development (Brand et al. 1996, Driever et al. 1996, Gaiano et al. 1996, Haffter et al. 1996, Jiang et al. 1996, Neuhauss et al. 1996, Piotrowski et al. 1996, Schilling et al. 1996a, b, Whitfield et al. 1996). No genes have been allocated to most of these mutations, but it holds great promise for the future.

IX 2 3 STUDYING OTHER CONGENITAL SYNDROMES PROVIDES INSIGHT INTO COMPLEX GENOTYPE-PHENOTYPE RELATIONSHIPS

Studies of the genetics of other diseases, syndromes and malformations, have shown that nature can have some interesting surprises for us in stall. In the introduction I discussed the *RET* proto-oncogene, a gene involved in no less than four different diseases. Loss-of-function mutations cause Hirschsprung's disease and various constitutive activation mutations cause MEN2A, MEN2B or FMTC cancers. The type (position) of mutation and the phenotype are strictly correlated. It is remarkable that one gene can cause such varied phenotypes. True, three of them are carcinomas, but they are different carcinomas nevertheless. Where *RET* is interesting for being associated with four different diseases, Hirschsprung's disease is interesting for being associated with five genes, *RET* among them (see introduction). Yet another interesting example of the muddled relationships between genes and diseases is Charcot-Marie-Tooth disease, a common peripheral nervous system affliction. Several loci, on chromosome 1, 17, X and possibly others, are connected to this inherited disorder (Patel and Lupski 1994). On chromosome 17 the *PMP22* gene is responsible, interestingly through two different mutational mechanisms: both a duplication of the gene and a dominant mutation can cause CMT1. Homozygous duplication causes a more severe phenotype, suggesting gene dosage effects. To make things more interesting, a mutation (different from the one for CMT1) in or hemizyosity of *PMP22* cause the neurological disorder

HNPP (Hereditary Neuropathy with liability to Pressure Palsies). Yet another mutation (a recessive one) in combination with hemizygosity causes CMT1. Dominant mutations of the *CX32/GJB1* gene on chromosome 1 and of the *MPZ* gene on chromosome X can cause CMT as well. Recapitulating we have one gene with three different types of mutation associated with two different diseases and one disease associated with at least three different genes, one of which can be mutated via two different mechanisms. Another genetically heterogeneous syndrome is Opitz syndrome, part of the CATCH22 family. Not only are loci on chromosome 22 linked to this syndrome, a chromosome Xp22 locus is associated with Opitz as well (Robin *et al.* 1995, Quaderi *et al.* 1997).

Examples like described above are important to bear in mind when trying to decide or identify which genes are causative for a syndrome. The possibility of long range effects of chromosomal aberrations on gene regulation as described earlier doesn't make this decision or identification any easier. Especially not if one considers the example of the Kreisler mouse. Extremely long range effects play no role here, but the effect of a chromosomal rearrangement on the expression of the candidate gene shows that one should take heed when studying phenotype-genotype relations. The Kreisler gene *MafB/Kr* is expressed in rhombomeres five and six and gives rise to the Kreisler mouse when deregulated due to an inversion event upstream of the gene (Cordes and Barsh 1994). The gene is actually located on the inverted segment, approximately 30 kb from the distal breakpoint. Homozygous *kr/kr*⁻ mice lack r5 and r6 and are affected in the development of the inner ear and the hyoid bone (Frohman *et al.* 1993, McKay *et al.* 1994). They show abnormal circling behavior (Kreisler is German for circler) and are deaf. *MafB/Kr* is expressed in the affected rhombomeres and their neural crest, explaining the causal relationship between the mutation and the phenotype. Eichmann *et al.* show that it is not that simple. In both chicken and mouse the expression of the gene extends beyond only these rhombomeres: it is found in differentiating neurons, the mesonephros, the perichondrium and the hemopoietic system (Eichmann *et al.* 1997). These expression sites persist in Kreisler

mice and there appear to be no defects in these systems. In other words, we are dealing with tissue specific gene inactivation due to chromosomal rearrangements and not with a simple on-off situation.

IX 4 ONGOING RESEARCH INTO HIRA REVEALS THAT IT IS ALMOST CERTAINLY INVOLVED IN CHROMATIN INTERACTIONS

A few years ago the HIR family consisted of only five members (two yeast and one each in human and worm), but the HIR family of proteins has been growing ever since, with a *Drosophila* (Dhh, *Drosophila* HIRA Homolog) and a *Fugu* homolog recently described (Kirov *et al.* 1998, Llievadot *et al.* 1998a, b) and an unpublished *Xenopus* version mentioned (Scamps *et al.*, unpublished). Our division of HIRA in three domains of relatively high, low and medium conservation (from N to C terminal) is confirmed by alignment of existing HIRs with *Fugu* HIRA and Dhh. The cloning of a *Drosophila* and *Xenopus* homolog brings new opportunities for modifying HIRA expression, since these two organisms are exceptionally amenable to reverse genetic analysis. *Fugu* HIRA could be very useful for mouse transgenic experiments, since the whole genomic size in *Fugu* is a mere 9 kb (*i.e.* ten percent of the human/mouse size).

Several recent papers already offer us some more insight into the HIRA network. Lorain *et al.* fished for HIRA interacting proteins with the two-hybrid system and describe one of them as a DnaJ family protein (Lorain *et al.* 1998). DnaJ is a part of the larger group of Dna heat shock proteins or chaperonins. Proteins in this family are involved in (re)folding of proteins, often in the context of multimolecular complexes. The new DnaJ member is called HIRIP4 (HIRA Interacting Protein 4) and is special in that it has a nuclear localization, as opposed to the cytoplasmic localization of previously described family members. Two putative zinc fingers in HIRIP4 are necessary and sufficient to bind HIRA, and HIRIP4 can bind to itself and to histone proteins. Since HIRA itself is able to interact with core histones H3.3B and H2B directly (Lorain *et al.* 1998), there seems to be ample evidence for our initial suggestion that HIRA is involved in chromatin interactions (Wilming *et al.* 1997). A link between HIRA and neural crest development is not only provided by the expression pattern described in this thesis, but also by recent findings by Magnaghi *et al.* Apparently Hira interacts with Pax3 and Pax7 (Magnaghi *et al.* 1998). In the introduction I described Pax3 and Pax7 knock-outs and that both result in neural crest related malformations. PAX3 hemizygosity is related to Waardenburg Syndrome, but many double knock-out mice die from a DGS like phenotype. HIRA binding to the Pax molecules may involve the Pax homeodomain; if this is confirmed it is a small step to suggest looking for interactions with other homeodomain proteins such as Hox and Msx. Association of HIRA with Pax transcription factors fuels the initial idea that HIRA is a transcription (co-)factor. A picture is starting to emerge where HIRA acts as a general transcriptional co-regulator recruited by specific transcription factors (*e.g.* PAX3, PAX7) and modulating chromatin. Haplo-insufficiency or absence of any or some parts of the multi-protein complex could bring this

mechanism to a halt, which could explain the many (partial) phenocopies of and overlaps with DGS. Preliminary results from *Hira* knock-out mice produced by our and other groups suggest some kind of perturbation of the cell-cycle. In vitro results show a shortened cell-cycle, enhancing neural crest cell growth and impeding differentiation. In vivo, $-/-$ knock-out embryos die *in utero* at E8.5-10 after apparently asynchronous/uncoordinated development of different parts of the embryo. Further study of the phenotype and its physiological or developmental background will be necessary to clarify the role and function of *Hira*.

IX 2.5 IS THE 'DIGEORGE GENE' STILL TO BE IDENTIFIED?





HIRA remains a prime candidate (see above), but given the reservations described earlier, it is very difficult to reject or put forward a CATCH22 gene *a priori*. Most of the 22q11 genes described in this thesis could conceivably be etiologic for CATCH22. After all, from the introduction it should be clear that there are many, in number and type, molecules guiding the proper development of neural crest and its derivatives. Anything from a cell-surface, trans-membrane or extra-cellular type to a kinase, transcription factor or receptor type protein is a potential candidate. GSCL is a good candidate since it is a transcriptional regulator and *Gooseoid* knock-out causes severe craniofacial deformities due to first and second arch malformation. A word of caution is in place: the homology between GSCL and *Gooseoid* is based on the homeodomain and a small N-terminal domain. Outside the homeodomain region are the sequences that one would assume confer the identity of the particular protein as different from other homeobox proteins with very similar homeodomains. In these sequences there is no significant homology. Moreover, *Gscd* knock-out mice are normal. Genes coding for transmembrane proteins like IDD and TRAINMODEL could play a role in regulating adhesion to ECM components or in signaling, adding to their candidate status. TRAINMODEL is highly expressed in the lung and heart and where we know much about the heart malformations as part of CATCH22, abnormal lung development is still uncharted territory but could be part as well (McDonald-McGinn *et al.* 1995). TOMCAT is likely to be involved in ligand induced signaling related to cell movement and adhesion, making it a prime candidate (I have discussed the role of catenins in the introduction). TBX1, as a transcriptional regulator, is another prime candidate with a fitting embryonic expression pattern and a family member associated with heart disease. A kinase protein like DGSG can be involved in signal transduction or any number of cellular processes and could thus play a pivotal role. The same could be said for RANBP1, probably involved in regulating nuclear processes. COMT is involved in neurotransmitter metabolism, so it can be responsible for ancillary components of the phenotype (*i.e.* psychotic illness), but is not responsible for the main components (neural crest, pharyngeal arch related). UFD1L, HUDDL, MORF and CLTCL1 appear on the surface to be involved in general cellular processes (protein degradation, cytokinesis, metabolism, and endocytosis, respectively), and GPIBB seems to be specifically active in platelets, so

they would not seem to be prime candidates. Interestingly, a report was published recently on UFD1L and its role in cardiac and craniofacial defects (Yamagishi *et al.* 1999). Yamagishi *et al.* started with *dHand* knock-out mice (*dHand* influences cardiac neural crest; see introduction) and analyzed which genes are downregulated as a result of the *dHand* null mutation. One of these genes is *Ufd1l*. Expression studies show that mRNAs can be found in embryonic structures relevant to DiGeorge syndrome (*i.e.* branchial arches, cardiac outflow tract, palate, frontonasal mass). In their patient series, all patients with known 22q11 deletions are hemizygous for *UFD1L*. This is not so surprising, since the gene is located in the commonly deleted region. Probably most of the deleted patients miss one copy of *HIRA* as well. More importantly, they found the first three *UFD1L* exons missing in one (out of ten) patient who was syndromic but did not have a FISH-detectable deletion. Preliminary results of similar experiments in European laboratories do at present show no deletions in these types of patients (Wadey *et al.* 1999). In conclusion, it is very likely that UFD1L plays an important role in 22DS etiology, but it is more than likely (and the opinion of probably the majority of researchers in the field) that other genes, either in or outside the DiGeorge region, have equally important parts to play. Before we can complete the picture describing 'Which?', 'Why?' and 'How?', we will have to perform much more extensive and more detailed expression, functional and mutational studies. Our search for the elusive DiGeorge gene(s) is not over yet.

IX 2.6 STUDYING CATCH22 PHENOCOPIES AND LIMB DEVELOPMENT MAY YIELD VALUABLE CLUES ABOUT THE REGULATORY PATHWAYS IN CRANIOFACIAL DEVELOPMENT

Earlier I indicated that Pax3 and HIRA are active in the same pathway, suggesting that this strengthens the case for HIRA. We may want to look at the best phenocopies of CATCH22, such as displayed by for example *Hoxa3* knock-outs, and carefully study the pathways of these genes in relation to 22q11 genes. In a sense this was how the role of UFD1L was discovered (see chapter IX-2-5):

Et1 knock-out mice are DGS phenocopies that display *dHand* downregulation (see **chapter II-4-I-2**) which in turn leads to *Ufdll* downregulation. Maybe more linked genes can be discovered this way, resulting in other strong candidate genes. Similarities between the development of limbs and the face in the context of 22q11 genes may also tell us more about the position of a gene as a candidate gene. Limb and face development differ significantly in important respects (the face doesn't have an Apical Ectodermal Ridge or Zone of Polarizing Activity), but on the genetic and developmental level there are some remarkable similarities. In the introduction we have already seen Greig Syndrome, a syndrome that combines craniofacial dysmorphism with limb malformations. There are more such syndromes (for example Acro-Oto-Ocular (Bertola *et al.* 1997), Crouzon (Murdoch-Kinch and Ward 1997), Cornelia de Lange (Ireland and Burn 1993, Allanson *et al.* 1997), Oral-Facial-Digital (OFD) (Dodge 1970, Whelan *et al.* 1975), Filippi (Orrico and Hayek 1997) and Smith-Lemli-Opitz (SLOS) (Lanoue *et al.* 1997)) and as mentioned earlier, CATCH22 patients sporadically present with limb abnormalities. In both systems epithelial-mesenchymal interactions play a crucial role and are interchangeable to a certain extent. Mesenchyme from either of the three facial primordia (maxillary, mandibular and frontonasal) can

-  population at large
-  undiagnosed 22DS patients
-  detected 22DS patients
-  patients falsely diagnosed as 22DS

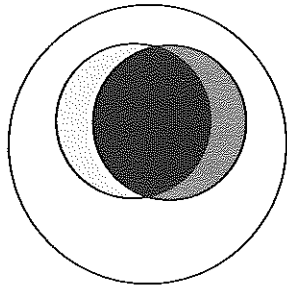


figure 36 A Venn diagram showing the real life situation faced by researchers and clinicians involved in phenotype based research. Not all 22DS patients are detected (light gray) and some non-CD225 patients are incorrectly diagnosed as 22DS (medium gray). Development of better screening tools together with increased clinical knowledge and training will bring the two circles closer together.

maintain limb AER and limb AER can induce chick frontonasal mesenchyme to form a reasonable facsimile of an upper beak. Maxillary or mandibular ectoderm in turn can support limb outgrowth. Results from transplantations of mesenchyme from hind limb bud (leg) to fore limb bud (wing) in chick embryos are comparable to transplantation of presumptive first arch neural crest to the position of second arch neural crest (Zwilling 1955, Saunders *et al.* 1959, Noden 1983a, Gumpel-Pinot 1984). In both situations the ectopic structures (bones, musculature) are reminiscent of those that would have been formed at the original site. Genetically, genes like *Msx1* and *Msx2* can be detected in both limbs and facial primordia, possibly working in a similar fashion. Other homeobox genes, from the *Hox* clusters, are key factors in patterning of the head (see introduction), but they set up separate gradients in the limb buds (Oliver *et al.* 1988, 1989, Wedden *et al.* 1989, Duboule 1992, Hunt and Izpisua-Belmonte and Duboule 1992, Krumlauf 1992, Tabin 1992, 1995, Johnson *et al.* 1994, Favier and Dolle 1997, Innis 1997, Rijli and Chambon 1997). Upstream regulators like RA affect both limb and craniofacial development (Denker *et al.* 1991, Wedden 1991, Mendelsohn *et al.* 1992, Noden 1992, Richman 1992, Morris-Kay 1993, Tabin 1995, Morris-Kay and Sokolova 1996, Tamura *et al.* 1997). In conclusion, it is clear that in many instances the same signalling pathways appear to be used in the face and the limb; studying limb development could therefore give us some insight into facial genetics.

IX 2 7 CLINICAL CONTRIBUTION TO OUR KNOWLEDGE IS EXPANDING AND IMPROVING

Since CATCH22 is detected based on phenotypical characteristics, identification of patients referred to as being syndromic is prone to human error. General practitioners are sometimes the first line of diagnosticians and they may not have the specialized knowledge to make the correct diagnosis, missing out on an unknown number of patients and probably also including patients actually not syndromic for CATCH22 (figure 36). Better information and education improves the situation, even though the spectrum of 22DS phenotypes is widening. Associations of new malformations with 22q11 deletions are found on a regular basis (Giannotti *et al.* 1994, Jaquez *et al.* 1997, Sabry *et al.* 1998), but in most cases this is based on very few patients, so some associations may be coincidental. But as our knowledge grows, we are able to study patients more carefully, picking up abnormalities that escaped our attention earlier and identifying patients we would not have recognized before. This, together with improvements in surgical techniques, could face us with the ironic effect of a rising incidence of CATCH22. Similar findings were done for congenital heart disease. Patients are diagnosed earlier than before and they are being subjected to corrective surgery when possible. Patients that would not have survived till reproductive age in the past, or would be socially unfit to reproduce if they did, can now transmit their mutated genes or chromosomes to a next generation. Growing knowledge of 22DS genetics, to which I hope this thesis has contributed, could help alleviate this side effect through better prenatal diagnostics and genetic counselling.

X

Materials & Methods

X 1 PROBES

NB84 is a 1.0 kb plasmid probe for locus D22S183, which has been assigned to the region 22pter-q11 (van Biezen *et al.* 1993) and was kindly provided by Dr E.C. Zwarthoff. It recognizes a *Pst*I polymorphism (Lekanne Deprez *et al.* 1991). Plasmid HP500 (Carey *et al.* 1992) and cosmids sc4.1 (Carey *et al.* 1992) and sc11.1 (Halford *et al.* 1993a) were kindly provided by Dr P.J. Scambler, and cos39 (Aubry *et al.* 1993) by Dr M. Aubry. A cosmid for the locus D22S75 (N25) (Driscoll *et al.* 1993) was purchased from Oncor (Gaithersburg, MD; URL4). This probe is supplied as digoxigenin-labelled DNA by the manufacturer and is premixed with a digoxigenin-labelled control cosmid (pH17) for the locus D22S39 in 22q13.3. Probes for the human β -globin gene and the Immunoglobulin- λ light chain complex were used as controls in Southern hybridizations.

C5 is a plasmid containing 2.8 kb of human *HIRA/TUPLE1* cDNA sequence (Halford *et al.* 1993c) and was kindly provided by Dr P.J. Scambler. Cosmid ch748 (48H7) was used as a *HIRA/TUPLE1* specific FISH probe (Halford *et al.* 1993c). Cosmids M69 (87D4) and M78 (114B1) were isolated using a plasmid containing a fragment of the PDGFB gene, which is located on 22q13. These were used as chromosome 22 specific control probes.

X 2 PATIENTS AND CELL LINES

Patients were ascertained through clinical geneticists. Blood samples or skin biopsies for the establishment of fibroblast cultures were obtained from probands and, in some cases, their parents. Cell lines GM03476, GM05878 and GM00980 were obtained from the NIGMS Human Mutant Cell Repository (Camden, NJ; URL5). GM03479 is from a DGS patient and has no detectable cytogenetic abnormality (Scambler *et al.* 1991). GM05878 is from the unaffected father of a DGS child and has a balanced translocation

t(10;22)(q26.3;q11.2) (Cannizzaro and Emanuel 1995). GM00980 is from a VCFS patient with a translocation t(11;22)(q25;q11) and monosomy 22pter-q11 (Fu *et al.* 1976). ADU is from a DGS patient with an apparently balanced translocation t(2;22)(q14.1;q11.1) (Augusseau *et al.* 1986) and was kindly provided by Dr F. Amblard. Hybrid cell lines were established by fusion of fibroblasts from patients with thymidine-kinase deficient Chinese hamster cells (A3) followed by selection of colonies in HAT culture medium according to standard procedures.

X 3 KARYOTYPING AND FISH

Cells were harvested from cultures of phytohaemagglutinin-stimulated lymphocytes, skin fibroblasts or established cell lines and spread onto slides for the production of G-banded or R-banded chromosome preparations and FISH. In some experiments, confluent fibroblasts were enriched for G1 interphase cells by growth in medium with 2% fetal calf serum for 4 days.

FISH was carried out as described by Arnoldus *et al.* (1990) with minor modifications. Cosmid probes were labelled by nick-translation with either biotin-16-dUTP or digoxigenin-11-dUTP. The hybridization mixture contained 5n ng of each labelled cosmid plus 1 mg human genomic DNA as a competitor in 10 ml 70% formamide, 2XSSC, (1X SSC = 150mM NaCl, 15 mM Na-citrate, pH7.0). After hybridization and washes, the slides were stained in three consecutive steps: (1) avidin-fluorescein, (2) biotinylated goat anti-avidin plus sheep anti-digoxigenin-rhodamin, (3) avidin-D-fluorescein plus donkey anti-sheep-Texas red. The slides were finally embedded in a glycerol mixture containing an anti-fade reagent and 4,6-diamino-2-phenyl-indole (DAPI) for counterstaining. In case of single color labelling (fluorescein), propium iodide was used as the counterstain. The slides were analyzed with a Zeiss Axioskop epifluorescence microscope, using various filter combinations (single, double and triple band).

For 22q11 deletion analysis, metaphase spreads or G1 interphase cells were hybridized simultaneously with the test probe and a chromosome 22q specific control probe (M69 or M78). The probes were usually stained in two colors, but single color staining was also used for metaphase preparations. Metaphases (at least 15 per case) in which both chromosomes 22 gave a signal with the control probe were scored for test probe signals in 22q11. Alternatively, interphase nuclei (at least 50 per case) with two signals for the control probe were scored for the number of test probe signals. In preparations of unaffected controls 95-100% of the metaphases and 90-95% of the interphase nuclei gave a double (dizygous) signal for the test probe, indicating high hybridization efficiency.

✕ 4 SOUTHERN ANALYSIS

✕ 4 1 HUMAN DNA

Genomic DNA, extracted from blood samples, cultured fibroblasts or established cell lines, was digested with restriction enzymes according to manufacturer's recommendations (Boehringer Mannheim, Mannheim, Germany; **URL6**) (New England Biolabs, Beverly, MA; **URL7**) (Pharmacia, Uppsala, Sweden; **URL8**) (Eurogentec, Seraing, Belgium; **URL9**), separated by agarose gel electrophoresis and transferred to HyBond N⁺ (Amersham, Little Chalfont, Bucks, UK; **URL10**) or Qiabrene⁺ (Qiagen, Hilden, Germany; **URL11**). DNA probes were labelled with [α -³²P]dATP and in some cases also with [α -³²P]dCTP, using the random primer labelling method.

Human placenta DNA was isolated from freshly frozen tissue and sonicated to a size range around 100 bp.

✕ 4 2 ZOO BLOT

A zoo blot containing EcoRI digested genomic DNA from human, Rhesus monkey, Sprague-Dawley rat, BALB/c mouse, dog, cow, rabbit, chicken and yeast *Saccharomyces cerevisiae*, was purchased from Clontech (Palo Alto, CA; **URL12**) and hybridized according to manufacturer's instructions (hybridization at 65 °C, wash in 2X SSC at 60 °C).

✕ 5 GRIDDED LIBRARY

The chromosome 22 specific cosmid library L22NCO3 used in this work was constructed at the Biomedical Sciences Division, Lawrence Livermore National Laboratories (Livermore, CA) under the auspices of the National Laboratory Gene Library project sponsored by the US Department of Energy (kindly provided by Dr P. de Jong). Cosmid containing bacteria from the frozen stock plates were plated out in a 6x6 array on HyBond N⁺ membranes with a BioMek robot (Beckmann, Fullerton, CA; **URL13**) using custom software (W. van Loon). Filters were subsequently treated according to the recommendations of the library producer. Filter hybridizations were preceded by pre-adsorption of vector and repeated sequences. Competitor DNA consists of a total of 500 μ g DNA: equal amounts of sheared human placenta DNA and cosmid DNA (*Sau3A* digested insert-less cosmids M100 or M89 or non-overlapping cosmids). Competitor DNA and radio-labelled cosmid probe were denatured and hybridized 2 hours at 65 °C in 5xSSC before being added to the filters.

✕ 6 SINGLE COPY PROBES

Restriction enzyme digested cosmids were exposed to Southern hybridization with 800 ng radio-labelled sheared human placenta DNA and radio-labelled non-overlapping cosmid DNA. Restriction bands failing to hybridize with both probes were considered single copy non-vector probes and gel-isolated and subcloned in a subsequent experiment. Prospective single copy probes were hybridized back to a Southern blot containing restriction enzyme digested human placenta DNA to ascertain the absence of non-unique sequences.

✕ 7 CONTIG ASSEMBLY

Sequences were retrieved from the EMBL databank (release 49) (AC000067-AC000071, AC000073-AC000080, AC000082-AC000097) and the linear order was established according to the annotation information. Overlaps were determined using the FastA program (Pearson and Lipman 1988). Contiguous sequences were assembled and were labeled 22q11 contig 1 through 5 from proximal to distal. The contigs are assembled from the following cosmids: contig 1: 103a2 (1-39862) + f41c7 (5065-43728) + dgcrca (8781-108400); contig 3: f39e1 (1-10909) + 72f8 (1-100090) + 98c4 (34-38718) + 49c12 (1880-39546) + 102g9 (15200-43944) + 83c5 (6268-42087); 129f8 (9363-43714) + 59f (15181-42027) + 105a (14400-38479) + 81h (2869-43485) + 31e (reverse complement 33656-1); contig 4: 100h (1-40649) + carlaa (1505-42672) + 91c (6643-44937) + 89h (854-39756) + 2h (9257-41098) + 33e (622-44426) + 8c (752-45539) + 56c (15227-44096) + 20b (10140-44456); contig 5: 68f (1-39579) + 24b (11184-38457) + p201m18 (reverse complement 139812-139806) + 68a1 (1-44575) + p201m18 (reverse complement 95242-1). Contig 2 is dgcrtel.

Consecutive 40 kb stretches of assembled sequences were used as query sequences for database comparison by Blast (Altschul *et al.* 1990) through the EMBL/Genbank, EST, translated yeast and translated fly databases. These databases are maintained by the Dutch CAOS/CAMM bioinformatics service in Nijmegen and available on their server (URL16). Prior to database searches the query sequences were masked with the Xblast program (Claverie 1994) to mask repeat sequences. Blast searches were performed with default parameters, with the exception of BlastX searches, where the low complexity filter was invoked. All programs run under the GCG (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics computer Group, 575 Science Drive, Madison, WI 53771, USA) and EGCG (Program Manual for the EGCG package, Peter Rice, the Sanger Center, Hinxton, Cambridgeshire, CB10 1SA, Great Britain) packages. Blast outputs were manually searched for significant homologies. Positive ESTs or proteins were compared with the relevant contig with FastA, TfastA or FrameAlign to establish the exact regions of homology. When ESTs were involved, paired ESTs (*i.e.* 5' and 3' ends of the same cDNA clone) were identified and orphan (unpaired) ESTs were checked for corresponding partners. If any were found, these were also mapped to the contigs by FastA. 100-200 kb Segments of contigs were reanalyzed and reannotated with the genomic analysis and annotation system in use at the Sanger Centre, which includes Blast and ACeDB (URL2) amongst a host of other dedicated software.

Contig sequences were analyzed by X-grail version 1.3c (Uberbacher and Mural 1991) in 50 kb segments, with 10 kb overlap. Searches for exons, P0III promoters, CpG islands and polyadenylation signals were performed using default parameters. X-grail was running on a Sun SPARC workstation.

Cosmids positive for NB84 were isolated from the Lawrence Livermore Laboratories sorted human chromosome 22 cosmid gridded library L22NC03 (constructed at the Biomedical Science Division, Lawrence Livermore National Laboratory, under the auspices of the National Laboratory Gene Library Project sponsored by the US Department of Energy and kindly provided by Dr Pieter de Jong). Cosmids were used as probes to isolate overlapping cosmids. Overlap was established and mapped using *EcoRI* restriction maps, cross-hybridizations and single copy probe hybridization. These single copy probes were isolated by hybridizing a Southern blot of *PstI* digests of several cosmids with labeled total human placenta DNA. Non-hybridizing restriction fragments were selected and used as single copy probes. Hybridizations were performed under standard conditions. YAC y20e1 is from an ICRF library constructed by Tony Monaco. PAC b20-280 is from a Lawrence Livermore library provided by Pieter de Jong (see above). YAC y30fh10 is from an ICI library.

Embryos were derived from crosses between FVB/N mice. The morning of vaginal plug was considered 0.5 days *post coitum* (dpc). The embryos were staged according to the number of developmental days *post coitum* (E8.5 - E12.5). The E8.5 embryos were staged under the dissecting microscope by counting somites.

X 12 CDNA CLONING, SEQUENCING AND DATABASE SEARCHES

The mouse embryo E11.5, the 26 week human fetal brain and the human testes cDNA libraries were from Clontech (Palo Alto, CA; [URL12](#)) (ML 1027a, HLI065b and HLI010b respectively). Another human fetal brain (17-18 weeks) cDNA library was purchased from Stratagene (La Jolla, CA; [URL14](#)) (#936206). Plaques were lifted onto nitrocellulose membranes (Millipore HATF filters) and hybridized as recommended by the manufacturers. The cDNA clones were sequenced by primer walking using the Pharmacia (Uppsala, Sweden; [URL8](#)) T7 sequencing kit. Sequences were assembled, aligned and compared using the GCG package at the CAOS/CAMM server in Nijmegen ([URL16](#)). Database searches were performed using Blast programs ([Altschul et al. 1990](#)). Multiple sequence alignments were edited manually with the GeneDoc program and printed with BoxShade.

X 13 WHOLE MOUNT IN SITU HYBRIDIZATION

Complementary (antisense) or non-complementary (sense) RNA probes were synthesized from linearized DNA templates using the DIG-UTP labelling kit (Boehringer Mannheim, Mannheim, Germany; [URL6](#)). In situ hybridization on whole mouse embryos (E8.5-E11.5) was performed according to the protocol of [Wilkinson \(1992\)](#), with posthybridization washes modified as follows: 50 minutes and 70 minutes in solution 1 at 70 °C; 20 minutes in solution 1 and 2 (50%/50%) at room temperature; two times 20 minutes solution 2 at room temperature; twice 30 minutes in solution 2 with RNase A at 37 °C; 10 minutes solution 2 at room temperature; twice 30 minutes solution 3 at 65 °C; twice 30 minutes TBST at room temperature. Adsorption of anti-digoxigenin antibodies to embryo powder was omitted. Embryos were either photographed on film and subsequently scanned, or images were captured electronically with an electronic camera. Images were processed using Adobe Photoshop software.

X 14 RNA ISOLATION

Adult organs or embryos were homogenized in GIT or LiCl Urea with a tissue homogenizer. Total RNA was purified by ultracentrifugation on a CsCl cushion or by LiCl precipitation and organic extractions. polyA⁺ RNA was isolated from total RNA with the PolyATtract mRNA isolation system III (Promega, Madison, WI; [URL15](#)).

X 15 NORTHERN ANALYSES

The mouse embryo and a human multiple tissue Northern blots were purchased from Clontech (Palo Alto, CA; [URL12](#)). Blots were hybridized as per manufacturer's protocol. The mouse multiple tissue Northern blot and human fetal Northern blot were produced by running 5 µg polyA⁺ RNA per lane on a denaturing formaldehyde agarose gel, blotting in 20x SSC onto Qiaabrene⁺ nylon membrane (Qiagen, Hilden, Germany; [URL11](#)), and UV cross-linking in a Stratalinker (Stratagene, La Jolla, CA; [URL14](#)). Mouse tissues were from female FVB mice. Human fetus was from an abortion on clinical indications of a 13 week embryo suffering from MLD.

X 16 RNASE PROTECTION ASSAY

Complementary (antisense) or non-complementary (sense) [α -³²P]-UTP labelled RNA probes were synthesized from linearized *Hira* cDNA templates, gel-eluted, hybridized to total RNA at 50 °C and subjected to RNase protection assay according to [Gilman \(1987\)](#).

X 17 RT-PCR DETECTION OF ALTERNATIVE HIRA mRNA

RT-PCR was performed on 5 μ g total RNA or 1 μ g polyA⁺ RNA (from E8.5, E9.5, E10.5 and E14.5 mouse embryos and adult mouse brain, spleen, liver and lung) with a mix of 300 ng each of oligo-dT and random hexamer primers according to established protocols. The RT reaction product was used for the following PCR amplification: one cycle of 2 minutes 94 °C, 2 minutes 50 °C, 2 minutes 72 °C; 30 cycles of 1 minute 94 °C, 1 minute 50 °C, 1.5 minutes 72 °C and one cycle of ten minutes at 72 °C. Primer sequences: p1 = ctggaaggttgatctgg, p2 = ggtattacagatggccctt, p3 = tcgcagactccaaatggtt, p4 = ccaatgtacgtagcccgtt, p5 = atactgctgct-gtgctgtg, p6 = ttgccaacgtaatgggg, p7 = cccagcaacagcatcactaa, p8 = atgctgagctgtctccta. Positive controls were 1 μ g of cDNAs I2B, E and F. Negative controls were omission of cDNA and a sample from an RT reaction without RNA.

data was collated and figures were drawn comparing the genomic structure of all the genes identified, with their mouse homologues. When the analysis and annotation was completed and the data base for mouse, the data for human, dog and rat were then used for large scale analysis and annotation.

XI

References

- Abo T, Balch CM**
1981 A differentiation antigen of human NK and K cells identified by a monoclonal antibody (HNK-1).
J Immunol **127**: 1024-1029
- Agrawal M, Brauer PR**
1996 Urokinase-type plasminogen activator regulates cranial neural crest cell migration in vitro.
Dev Dyn **207**: 281-290
- Ahn J, Ludecke HJ, Lindow S, Horton WA, Lee B, Wagner MJ, Horsthemke B, Wells DE**
1995 Cloning of the putative tumour suppressor gene for hereditary multiple exostoses (EXT1).
Nat Genet **11**: 137-143
- Aisenberg AC**
1984 New genetics of Burkitt's lymphoma and other non-Hodgkin's lymphomas.
Am J Med **77**: 1083-1090
- Akimaru H, Chen Y, Dal P, Hou DX, Nonaka M, Smolik SM, Armstrong S, Goodman RH, Ishii S**
1997 Drosophila CBP is a co-activator of cubitus interruptus in hedgehog signalling.
Nature **386**: 735-738
- Akimenko MA, Ekker M, Wegner J, Lin W, Westerfield M**
1994 Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head.
Neurosci **14**: 3475-3486
- Alfandari D, Wolfsberg TG, White JM, DeSimone DW**
1997 ADAM 13: a novel ADAM expressed in somitic mesoderm and neural crest cells during *Xenopus laevis* development.
Dev Biol **182**: 314-330
- Allanson JE, Hennekam RC, Ireland M**
1997 De Lange syndrome: subjective and objective comparison of the classical and mild phenotypes.
J Med Genet **34**: 645-650
- Altschul SF, Gish W, Miller W, Meyers EW, Lipman DJ**
1990 Basic local alignment search tool.
J Mol Biol **215**: 403-410
- Amati F, Mari A, Digilio MC, Mingarelli R, Marino B, Giannotti A, Novelli G, Dallapiccola B**
1995 22q11 deletions in isolated and syndromic patients with tetralogy of Fallot.
Hum genet **95**: 479-482
- Ammann AJ, Wara DW, Cowan MJ, Barrett DJ, Stiehm ER**
1982 The DiGeorge syndrome and the fetal alcohol syndrome.
Am J Dis Child **136**: 906-908
- Anderson DJ**
1989 The neural crest cell lineage problem: neurogenesis?
Neuron **3**: 1-12
- Anderson DJ**
1997 Cellular and molecular biology of neural crest cell lineage determination.
Trends Genet **13**: 276-280
- Andres G**
1946 Über Induktion und Entwicklung von Kopforganen aus Unkenknotodem im Molch (Epidermis, Plakoden und Derivate der Nemaaläste)
Rev Suisse Zool **53**: 502-510
- Andres G**
1949 Untersuchungen an Chiridren von Triton und Bombinator.
Genetics **24**: 387-534
- Angrist M, Bolk S, Halushka M, Lapchak PA, Chakravarti A**
1996 Germ-line mutations in glial cell line-derived neurotrophic factor (GDNF) and RET in a Hirschsprung disease patient.
Nat Genet **14**: 341-344
- Antoniades HN**
1991 PDGF: a multifunctional growth factor.
Baillieres Clin Endocrinol Metab **5**: 595-613
- Antoniades K, Peonidis A, Pehlivanidis C, Kavadia S, Panagiotidis P**
1994 Craniofacial manifestations of Smith-Lemli-Opitz syndrome: case report.
Int J Oral Maxillofac Surg **23**: 363-365
- Arai H, Hori S, Aramori J, Ohkubo H, Nakanishi S**
1990 Cloning and expression of a cDNA encoding an endothelin receptor.
Nature **348**: 730-732
- Arai E, Ikeuchi T, Nakamura Y**
1994 Characterization of the translocation break-point on chromosome 22q11.2 in a patient with neurofibromatosis type 2 (NF2).
Hum Mol Genet **3**: 937-939
- Arnoldus EPJ, Wiegant J, Noordermeer IA, Wessels JW, Beverstock GC, Grosveld GC, van der Ploeg M, Raap AK**
1990 Detection of the Philadelphia chromosome in interphase nuclei.
Cytogenet Cell Genet **54**: 108-111
- Arvystas M, Shprintzen RJ**
1984 Craniofacial morphology in the Velo-Cardio-Facial Syndrome
J Craniofacial Genet **4**: 39-45
- Antoniades HN**
1991 PDGF: a multifunctional growth factor.
Baillieres Clin Endocrinol Metab **5**: 595-613
- Asamoto K, Nojyo Y, Aoyama H**
1995 Restriction of the fate of early migrating trunk neural crest in gangliogenesis of avian embryos.
Int J Dev Biol **39**: 975-984
- Aubry M, Marineau C, Zhang FR, Zahed L, Figlewicz D, Delattre O, Thomas G, de Jong PJ, Julien J-P, Rouleau GA**
1992 Cloning of six new genes with zinc finger motifs mapping to short and long arms of human acrocentric chromosome 22 (p and q11.2).
Genomics **13**: 641-648
- Aubry M, Demczuk S, Desmaze C, Aikem M, Aurlas A, Julien J-P, Rouleau GA**
1993 Isolation of a zinc-finger gene consistently deleted in DiGeorge syndrome.
Hum Mol Genet **2**: 1583-1587
- Auerbach R**
1960 Morphogenetic interactions in the development of the mouse thymus gland.
Dev Biol **2**: 271-284
- Augusseau S, Jouk S, Jalbert P, Prieur M**
1986 DiGeorge syndrome and 22q11 arrangements.
Hum Genet **74**: 206
- Augustine KA, Liu ET, Sadler TW**
1995 Antisense inhibition of engrailed genes in mouse embryos reveals roles for these genes in craniofacial and neural tube development.
Teratology **51**: 300-310
- Axton R, Hanson I, Danes S, Sellar G, van Heyningen V, Prosser J**
1997 The incidence of PAX6 mutation in patients with simple aniridia: an evaluation of mutation detection in 12 cases.
J Med Genet **34**: 279-286
- Baker CV, Bronner-Fraser M, Le Douarin NM, Teillet MA**
1997 Early- and late-migrating cranial neural crest cell populations have equivalent developmental potential in vivo.
Development **124**: 3077-3087
- Baldini A, Jurecic V, Botta A, Cheah Y-C, Rivera S, Lindsay EA**
1998 Functional analysis of the DiGeorge syndrome region in mouse: generation and germ line transmission of chromosome deletions, duplications and single gene disruptions
Am J Hum Genet (Suppl) **63**: A114

- Ballabio A, Bardoni B, Carrozzo R, Andria G, Bick D, Campbell L, Hamel B, Ferguson-Smith MA, Gimelli G, Fraccaro M, et al**
1989 Contiguous gene syndromes due to deletions in the distal short arm of the human X chromosome.
Proc Natl Acad Sci U S A **86**: 10001-10005
- Ballabio A, Andria G**
1992 Deletions and translocations involving the distal short arm of the human X chromosome: review and hypotheses.
Hum Mol Genet **1**: 221-227
- Baldwin CT, Hoth CF, Macina RA, Milunsky A**
1995 Mutations in PAX3 that cause Waardenburg syndrome type 1: ten new mutations and review of the literature.
Am J Med Genet **58**: 115-122
- Barald K**
1988 Monoclonal antibodies made to chick mesencephalic neural crest cells and to olfactory ganglion neurons identify a common antigen on the neurons and a neural crest subpopulation.
J Neurosci Res **21**: 107-118
- Barbu M, Ziller C, Rog PM, LeDouarin NM**
1986 Heterogeneity in migrating neural crest cells revealed by a monoclonal antibody.
J Neurosci **6**: 2215-2225
- Bard JBL, Baldock RA, Davidson DR**
1988 Elucidating the genetic networks of development: a bioinformatics approach.
Genome Res **8**: 859-863
- Barroffo A, Dupin E, LeDouarin NM**
1988 Clone-forming ability and differentiation potential of migratory neural crest cells.
Proc Natl Acad Sci USA **85**: 5325-5329
- Barroffo A, Dupin E, LeDouarin NM**
1991 Common precursors for neural and mesectodermal derivatives in the cephalic neural crest.
Development **112**: 301-305
- Bassett AS, Hodgkinson K, Chow EW, Correlá S, Scutt LE, Weksberg R**
1998 22q11 deletion syndrome in adults with schizophrenia.
Am J Med Genet **81**: 328-337
- Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soultis J, Grayzel D, Kroumpouzou E, Trallis TA, Leblanc-Draceski J, Renault B, Kucherlapati R, Seidman JG, Seidman CE**
1997 Mutations in human cause limb and cardiac malformations in Holt-Oram syndrome.
Nature Genet **15**: 30-35
- Baynash AG, Hosoda K, Giald A, Richardson JA, Emoto N, Hammer RE, Yanagisawa M**
1994 Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons.
Cell **79**: 1277-1285
- Beauvais-Jouneau A, Delouève A, Craig SE, Humphries MJ, Thiery JP, Dufour S**
1997 Direct role of the carboxy-terminal cell-binding domain of fibronectin in neural crest cell motility.
Exp Cell Res **233**: 1-10
- Becker N, Seitanidou T, Murphy P, Mattei MG, Topilko P, Nieto MA, Wilkinson DG, Charnay P, Giliardi-Hebenstreit P**
1994 Several receptor tyrosine kinase genes of the Eph family are segmentally expressed in the developing hindbrain.
Mech Dev **47**: 3-17
- Bee J, Thorogood P**
1980 The role of tissue interactions in the skeletogenic differentiation of avian neural crest cells.
Dev Biol **78**: 47-62
- Belohradsky BH**
1985 Thyromastoplasie und -hypoplasie mit Hypoparathyreoidismus, Herz- und Gefäßstörungen (Di-George-Syndrom).
Ergeb Inn Med Kinderheilkd **54**: 35-105
- Bershof JF, Guyuron B, Olsen MM**
1992 G syndrome: a review of the literature and a case report.
J Craniomaxillofac Surg **20**: 24-27
- Bertola DR, Wolf LM, Toriello HV, Netzloff ML**
1997 Auro-oto-ocular syndrome: further evidence for a new autosomal recessive disorder.
Am J Med Genet **73**: 442-446
- Bidaud C, Salomon R, Van Camp G, Pelet A, Attie T, Eng C, Bonduelle M, Amiel J, Nihoul-Fekete C, Willems PJ, Munnich A, Lyonnet S**
1997 Endothelin-3 gene mutations in isolated and syndromic Hirschsprung disease.
Eur J Hum Genet **5**: 247-251
- Bielec PE, Gallagher DS, Womack JE, Busbee DL**
1998 Homologies between human and dolphin chromosomes detected by chromosome painting.
Cytogenet Cell Genet **81**: 18-25
- Birgbauer E, Fraser SE**
1994 Violation of cell lineage restriction compartments in the chick hindbrain.
Development **120**: 1347-1356
- Bischoff FR, Krebber H, Smirnova E, Dong W, Pongsigl H**
1995 Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1.
EMBO J **14**: 705-715
- Blinder MA, Marasa JC, Reynolds CH, Deaven LL, Tollefsen DM**
1988 Heparin cofactor II: cDNA sequence, chromosome localization, restriction fragment length polymorphism, and expression in *Escherichia coli*.
Biochemistry **27**: 752-759
- Blouin JL, Duriaux-Sail G, Chen H, Gos A, Morris MA, Rossier C, Antonarakis SE**
1996 Mapping of the gene for the p60 subunit of the human chromatin assembly factor (CAF1A) to the Down syndrome region of chromosome 21.
Genomics **33**: 309-312
- Bockman DE, Kirby ML**
1984 Dependence of thymus development on derivatives of the neural crest.
Science **223**: 498-500
- Bockman DE, Kirby ML**
1985 Neural crest interactions in the development of the immune system.
J Immunol **135**: 766S-768S
- Bockman DE, Redmond ME, Waldo K, Davis H, Kirby ML**
1987 Effect of neural crest ablation on development of the heart and arch arteries in the chick.
Am J Anat **180**: 332-341
- Bockman DE, Redmond ME, Kirby ML**
1989 Alteration of early vascular development after ablation of cranial neural crest.
Anat Rec **225**: 209-217
- Bolande RP**
1974 The neurocristopathies. A unifying concept of disease arising in neural crest maldevelopment.
Hum Pathol **5**: 409-429
- Bolande RP**
1981 Neurofibromatosis - the quintessential neurocristopathy: pathogenetic concepts and relationships.
Adv Neurol **29**: 67-75
- Bollag RJ, Siegfried Z, Cebra-Thomas JA, Garvey N, Davison EM, Silver L**
1994 An ancient family of embryonically expressed mouse genes sharing a conserved protein motif with the T-loxus.
Nature Genet **7**: 383-389
- Botta A, Lindsay EA, Jurecic V, Baldini A**
1997 Comparative mapping of the DiGeorge syndrome region in mouse shows inconsistent gene order and differential degree of gene conservation.
Mamm Genome **8**: 890-895

- Boucaut JC, Darriber T, Poole TJ, Aoyama H, Yamada KM, Thiery JP**
1984 Biologically active synthetic peptides as probes of embryonic development: a competitive peptide inhibitor of fibronectin function inhibits gastrulation in amphibian embryos and neural crest cell migration in avian embryos.
J Cell Biol **99**: 1822-1830
- Brahe C, Bannetta P, Meera Khan P, Arwert F, Serra A**
1986 Assignment of the catechol-O-methyltransferase gene to human chromosome 22 in somatic cell hybrids.
Hum Genet **74**: 230-234
- Brand M, Heisenberg CP, Warga RM, Pelegri F, Karlstrom RO, Beuchle D, Picker A, Jiang YJ, Furutani-Seiki M, van Eeden FJ, Granato M, Hafter P, Hammerschmidt M, Kane DA, Kelsh RN, Mullins MC, Odenthal J, Nusslein-Volhard C**
1996 Mutations affecting development of the midline and general body shape during zebrafish embryogenesis.
Development **123**: 129-142
- Brand M, Le Moullec JM, Corvol P, Gasc JM**
1998 Ontogeny of endothelins-1 and -3, their receptors, and endothelin converting enzyme-1 in the early human embryo.
J Clin Invest **101**: 549-559
- Brandli AW, Kirschner MW**
1995 Molecular cloning of tyrosine kinases in the early *Xenopus* embryo: identification of Eck-related genes expressed in cranial neural crest cells of the second (hyoid) arch.
Dev Dyn **203**: 119-140
- Brannan CJ, Perkins AS, Vogel KS, Ratner N, Nordlund ML, Reid SW, Buchberg AM, Jenkins NA, Parada LF, Copeland NG**
1994 Targeted disruption of the neurofibromatosis type-1 gene leads to developmental abnormalities in heart and various neural crest-derived tissues.
Genes Dev **8**: 1019-1029
- Bressan A, Somma MP, Lewis J, Santolamazza C, Copeland NG, Gilbert DJ, Jenkins NA, Lavia P**
1991 Characterization of the opposite strand genes from the mouse bidirectionally transcribed HTF9 locus.
Gene **103**: 201-209
- Breuning MH, Dauwerse HG, Fugazza G, Saris JJ, Spruit L, Wijnen H, Tommerup N, van der Hagen CB, Imazumi K, Kuroki Y, et al**
1993 Rubinstein-Taybi syndrome caused by submicroscopic deletions within 16p13.3.
Am J Hum Genet **52**: 249-254
- Briehl MM, Miesfeld RL**
1991 Isolation and characterization of transcripts induced by androgen withdrawal and apoptotic cell death in the rat ventral prostate.
Mol Endocr **5**: 1381-1388
- Bronner-Fraser M**
1985 Alterations in neural crest migration by a monoclonal antibody that affects cell adhesion.
J Cell Biol **101**: 610-617
- Bronner-Fraser M**
1986 An antibody to a receptor for fibronectin and laminin perturbs cranial neural crest development in vivo.
Dev Biol **117**: 528-536
- Bronner-Fraser M**
1988a Distribution and function of tenascin during cranial neural crest development in the chick.
J Neurosci Res **21**: 135-147
- Bronner-Fraser M**
1988b Inductive interactions underlie neural crest formation.
Adv Pharmacol **42**: 883-887
- Bronner-Fraser M**
1995 Origin of the avian neural crest.
Stem Cells (Dayt) **13**: 640-646
- Bronner-Fraser M, Lallier T**
1988 A monoclonal antibody against a laminin-heparan sulfate proteoglycan complex perturbs cranial neural crest migration in vivo.
J Cell Biol **106**: 1321-1330
- Bronner-Fraser M, Steber-Blum M, Cohen AM**
1980 Clonal analysis of the avian neural crest: migration and maturation of mixed neural crest clones injected into host chicken embryos.
J Comp Neurol **193**: 423-434
- Bronstein JM, Kozak CA, Chen XH, Wu S, Danciger M, Korenberg JR, Farber DB**
1995 Chromosomal localization of murine and human oligodendrocyte-specific protein genes.
Genomics **34**: 255-257
- Bronstein JM, Popper P, Micevych PE, Farber DB**
1996 Isolation and characterization of a novel oligodendrocyte-specific protein.
Neurology **47**: 772-778
- Bucan M, Gatalica B, Nolan P, Chung A, Leroux A, Grossman MH, Nadeau JH, Emanuel BS, Budarf M**
1993 Comparative mapping of 9 human chromosome 22q loci in the laboratory mouse.
Hum Mol Genet **2**: 1245-1252
- Budarf ML, Emanuel BS**
1997 Progress in the autosomal segmental aneuploidy syndromes (SAS): single or multi-locus disorders?
Hum Mol Genet **6**: 1657-1665
- Budarf M, Sellinger B, Griffin C, Emanuel BS**
1989 Comparative mapping of the constitutional and tumor-associated 11;22 translocations.
Am J Hum Genet **45**: 128-139
- Budarf ML, Collins J, Gong W, Roe B, Wang Z, Bailey LC, Sellinger B, Michaud D, Driscoll DA, Emanuel BS**
1995 Cloning a balanced translocation associated with DiGeorge syndrome and identification of a disrupted candidate gene.
Nature Genet **10**: 269-277
- Budarf ML, Eckman B, Michaud D, McDonald T, Gavigan S, Buetow KH, Tatsumura Y, Liu Z, Hilliard C, Driscoll D, Goldmuntz E, Meese E, Zwarthoff EC, Williams S, McDermid H, Dumanski JR, Biegel J, Bell CJ, Emanuel B**
1996 Regional localization of over 300 loci on human chromosome 22 using somatic cell hybrid panel mapping.
Genomics **35**: 275-288
- Budarf M, McDonald T, Sellinger B, Kozak C, Graham C, Wistow G**
1997 Localization of the human gene for macrophage migration inhibitory factor (MIF) to chromosome 22q11.2.
Genomics **39**: 235-236
- Burn J**
1989 The face and immune system in Tetralogy of Fallot.
Int J Cardiol **22**: 237-239
- Burn J, Takao A, Wilson D, Cross I, Momma K, Wade R, Scambler P, Goodship J**
1993 Conotruncal anomaly face syndrome is associated with a deletion within chromosome 22q11.
J Med Genet **30**: 822-824
- Buxton PG, Kostakopoulou K, Brickell P, Thorogood P, Ferretti P**
1997 Expression of the transcription factor slug correlates with growth of the limb bud and is regulated by FGF-4 and retinoic acid.
Int J Dev Biol **41**: 559-568
- Cannizzaro LA, Emanuel BS**
1985 In situ hybridization and breakpoint mapping III. DiGeorge syndrome with partial monosomy of chromosome 22.
Cytogenet Cell Genet **39**: 179-183
- Carey AH, Roach S, Williamson R, Dumanski JR, Nordenskjold M, Collins VP, Rouleau G, Blin N, Jalbert P, Scambler PJ**
1990 Localization of 27 DNA markers to the region of human chromosome 22q11-pter deleted in patients with the DiGeorge syndrome and duplicated in the der22 syndrome.
Genomics **7**: 299-306

- Carey AH, Kelly D, Halford S, Wadey R, Wilson D, Goodship J, Burn J, Paul T, Sharkey A, Dumanski J, Nordenskjold M, Williamson R, Scambler PJ**
1992 Molecular genetic study of the frequency of monosomy 22q11 in DiGeorge syndrome.
Am J Hum Genet **51**: 964-970
- Carlson C, Papolos D, Pandita RK, Faedda GL, Veit S, Goldberg R, Shprintzen R, Kucherlapati R, Morrow B**
1997a Molecular analysis of velo-cardio-facial syndrome patients with psychiatric disorders.
Am J Hum Genet **60**: 851-859
- Carlson C, Sirotkin H, Pandita R, Goldberg R, McKie J, Eadey R, Patanjali SR, Weissman SM, Anyane-Yeboah K, Warburton D, Scambler P, Shprintzen R, Kucherlapati R, Morrow BE**
1997b Molecular definition of 22q11 deletions in 151 velo-cardio-facial syndrome patients.
Am J Hum Genet **61**: 620-629
- Chapman DL, Agulnik I, Hancock S, Silver LM, Papaioannou YE**
1996a Tbx6, a mouse T-box gene implicated in paraxial mesoderm formation at gastrulation.
Dev Biol **180**: 534-542
- Chapman DL, Garvey N, Hancock S, Alexiou M, Agulnik SI, Gibson-Brown JJ, Cebra-Thomas J, Bollag RJ, Silver LM, Papaioannou YE**
1996b Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development.
Dev Dyn: 206: 379-390
- Chareonvit S, Osumi-Yamashita N, Ikeda M, Eto K**
1997 Murine forebrain and midbrain crest cells generate different characteristic derivatives in vitro.
Dev Growth Differ **39**: 493-503
- Chazaud C, Oulad-Abdelghani M, Bouillet P, Decimo D, Chambon P, Dolle P**
1996 AP-2.2, a novel gene related to AP-2, is expressed in the forebrain, limbs and face during mouse embryogenesis.
Mech Dev **54**: 83-94
- Chieffo C, Garvey N, Gong W, Roe B, Zhang G, Silver L, Emanuel BS, Budarf M**
1997 Isolation and characterization of a gene from the DiGeorge Chromosomal Region homologous to the mouse Tbx1 gene.
Genomics **43**: 267-277
- Chin JR, Werb Z**
1997 Matrix metalloproteinases regulate morphogenesis, migration and remodeling of epithelium, tongue skeletal muscle and cartilage in the mandibular arch.
Development **124**: 1519-1530
- Choi DS, Ward SJ, Messaddeq N, Launay JM, Maroteaux L**
1997 5-HT2B receptor-mediated serotonin morphogenetic functions in mouse cranial neural crest and myocardial cells.
Development **124**: 1745-1755
- Chowdhary BP, Fronicke L, Gustavsson I, Scherthan H**
1996 Comparative analysis of the cattle and human genomes: detection of ZOO-FISH and gene mapping-based chromosomal homologies.
Mamm Genome **7**: 297-302
- Ciment G, Weston JA**
1985 Segregation of developmental abilities in neural crest-derived cells: identification of partially restricted intermediate cell types in the branchial arches of avian embryos.
Dev Biol **111**: 73-83
- Claverie J-M**
1994 Large scale sequence analysis. In: Adams MD, Fields C, Venter JC (Eds)
Automated DNA sequencing and analysis techniques. Academic, New York: 267-279
- Clouthier D, Hosoda K, Richardson DA, Williams SC, Yanagisawa H, Kuwaki T, Kumada M, Hammer RE, Yanagisawa M**
1998 Cranial and cardiac neural crest defects in endothelin-A receptor-deficient mice.
Development **125**: 813-824
- Collazo A, Bronner-Fraser M, Fraser SE**
1993 Vital dye labeling of *Xenopus laevis* trunk neural crest reveals multipotency and novel pathways of migration.
Development **118**: 363-376
- Collins JE, Cole CG, Smink LJ, Garrett CL, Leversha MA, Soderlund CA, Maslen G, Everett LA, Rice KM, Coffey AJ, Gregory SG, William R, Dunham A, Davies AF, Hassock S, Todd CM, Lehrach H, Hulsebos TJM, Weissenbach J, Morrow B, Kucherlapati RS, Wadey R, Scambler PJ, Kim UJ, Simon MI, Peyrard M, Xie YG, Carter NP, Durbin R, Dumanski JP, Bentley DR, Dunham I**
1995 A high-density YAC contig map of human chromosome 22.
Nature **377** [Suppl]: 367-379
- Colombo P, Yon J, Fried M**
1991 The organization and expression of the human U7a ribosomal protein gene.
Biochim Biophys Acta **1129**: 93-95
- Colombo P, Yon J, Garson K, Fried M**
1992 Conservation of the organization of five tightly clustered genes over 600 million years of divergent evolution.
Proc Natl Acad Sci USA **89**: 6358-6362
- Conley ME, Beckwith JB, Mancer JFK, Tenckhoff L**
1979 The spectrum of the DiGeorge syndrome.
J Pediatr **94**: 883-890
- Consigliere S, Stanyon R, Koehler U, Agoramoorthy G, Wienberg J**
1996 Chromosome painting defines genomic rearrangements between redhowler monkey subspecies.
Chrom Res **4**: 264-270
- Conway SJ, Godt RE, Hatcher CJ, Leatherbury L, Zolotouchnikov VY, Brotto MAP, Copp AJ, Kirby ML, Creazzo TL**
1997a Neural crest is involved in development of abnormal myocardial function.
J Mol Cell Cardiol **29**: 2675-2685
- Conway SJ, Henderson DJ, Copp AJ**
1997b Pax3 is required for cardiac neural crest migration in the mouse: evidence from the *spotch* (Sp2H) mutant.
Development **124**: 505-514
- Corbo JC, Erives A, Di Gregorio A, Chang A, Levine M**
1997 Dorsoventral patterning of the vertebrate neural tube is conserved in a protochordate.
Development **124**: 2335-2344
- Cordes SP, Barsh GS**
1994 The mouse segmentation gene *kr* encodes a novel basic domain-leucine zipper transcription factor.
Cell **79**: 1025-1034
- Couly G**
1981 Les neurocristopathies du bourgeon naso-frontal humain. Les syndromes ethmoïdés (hypo et hyperseptothoïdismes).
Rev Stomatol Chir Maxillofac **82**: 213-225
- Couly G, Le Douarin NM**
1990 Head morphogenesis in embryonic avian chimeras: evidence for a segmental pattern in the ectoderm corresponding to the neurocrest.
Development **108**: 543-558
- Couly GF, Coltey PM, Le Douarin NM**
1993 The triple origin of skull in higher vertebrates: a study in quail-chick chimeras.
Development **117**: 409-429
- Couly G, Grapin-Botton A, Coltey P, Le Douarin NM**
1996 The regeneration of the cephalic neural crest, a problem revisited: the regenerating cells originate from the contralateral or from the anterior and posterior neural fold.
Development **122**: 3393-3407
- Couly G, Grapin-Botton A, Coltey P, Ruhin B, Le Douarin NM**
1998 Determination of the identity of the derivatives of the cephalic neural crest: incompatibility between Hox gene expression and lower jaw development.
Development **125**: 3445-3459

- Goh K, Yang J, Hynes R**
1997 Mesodermal defects and cranial neural crest apoptosis in (alpha)5 integrin-null embryos.
Development **124**: 4309-4319
- Goldberg R, Marion R, Borderon M, Wiznia A, Shprintzen RJ**
1985 Phenotypic overlap between velo-cardio-facial syndrome (VCF) and the D:George sequence (DGS).
Am J Hum Genet (Suppl): A54
- Goldberg R, Motzkin B, Marion R, Scambler PJ, Shprintzen RJ**
1993 Velo-cardio-facial syndrome: a review of 120 patients.
Am J Med Genet **45**: 313-319
- Goldmuntz E, Driscoll D, Budarf ML, Zackal EH, McDonald-McGinn DM, Biegel JA, Emanuel BS**
1993 Microdeletions of chromosomal region 22q11 in patients with congenital conotruncal cardiac defects.
J Med Genet **30**: 807-812
- Goldmuntz E, Wang Z, Roe BA, Budarf ML**
1996 Cloning, genomic organization, and chromosomal localization of human citrate transport protein to the D:George/Velocardiofacial syndrome minimal critical region.
Genomics **33**: 271-276
- Goldstein RS, Kalchelm C**
1991 Normal segmentation and size of the primary sympathetic ganglia depend upon the alternation of rostrocaudal properties of the somites.
Development **112**: 327-334
- Gong W, Emanuel BS, Collins J, Kim DH, Wang Z, Chen F, Zhang G, Roe B, Budarf ML**
1996 A transcriptional map of the D:George and velo-cardio-facial syndrome minimal critical region on 22q11.
Hum Mol Genet **5**: 789-800
- Gong W, Emanuel BS, Galili N, Kim NH, Roe B, Driscoll DA, Budarf ML**
1997 Structural and mutational analysis of a conserved gene (DGS1) from the minimal D:George syndrome critical region.
Hum Mol Genet **6**: 267-276
- Goodship J, Cross I, Scambler P, Burn J**
1995 Monozygotic twins with chromosome 22q11 deletion and discordant phenotype.
J Med Genet **32**: 746-748
- Gorlin RJ, Toriello HV, Cohen MM (Eds)**
1995 Hereditary hearing loss and its syndromes.
Oxford monographs on medical genetics no. 28. Oxford University Press, New York
- Gottlieb S, Emanuel B, Driscoll DA, Sellinger B, Wang Z, Roe B and Budarf M**
1997 The D:George syndrome critical region contains a Goosecoid-like (GSCL) homeobox gene that is expressed early in human development.
Am J Hum Genet **60**: 1194-1201
- Gottlieb S, Driscoll DA, Punnett HH, Sellinger B, Emanuel BS, Budarf ML**
1998 Characterization of 10p deletions suggests two nonoverlapping regions contribute to the D:George syndrome phenotype.
Am J Hum Genet **62**: 495-498
- Goureau A, Yerte M, Schmitz A, Riquet J, Milan D, Pinton P, Frelat G, Gellin J**
1996 Human and porcine correspondence of chromosome segments using bidirectional chromosome painting.
Genomics **36**: 252-262
- Graham A, Lumsden A**
1996 Patterning the cranial neural crest.
Biochem Soc Symp **62**: 77-83
- Graham A, Heyman I, Lumsden A**
1993 Even-numbered rhombomeres control the apoptotic elimination of neural crest cells from odd-numbered rhombomeres in the chick hindbrain.
Development **119**: 233-245
- Graham A, Francis-West P, Brickell P, Lumsden A**
1994 The signalling molecule BMP4 mediates apoptosis in the rhombencephalic neural crest.
Nature **372**: 684-686
- Graham A, Koentges G, Lumsden A**
1996 Neural crest apoptosis and the establishment of craniofacial pattern: an honorable death.
Mol Cell Neurosci **8**: 76-83
- Graveson AC, Smith MM, Hall BK**
1997 Neural crest potential for tooth development in a urodele amphibian: developmental and evolutionary significance.
Dev Biol **188**: 34-42
- Gray H, Williams PL (Eds)**
1995
Gray's Anatomy 38th Edition. Churchill Livingstone, New York
- Greenberg F, Courtney KB, Wessels RA, Huhta J, Carpenter RJ, Rich DC, Ledbetter DH**
1988a Prenatal diagnosis of deletion 17p13 associated with D:George anomaly.
Am J Med Genet **31**: 1-4
- Greenberg F, Elder FFB, Haffner P, Northrup H, Ledbetter DH**
1988b Cytogenetic findings in a prospective series of patients with D:George anomaly.
Am J Hum Genet **43**: 605-611
- Griffin CA, McKeon C, Israel MA, Geggone A, Ghysdael J, Stehelin D, Douglass EC, Green AE, Emanuel BS**
1986 Comparison of constitutional and tumor-associated 11;22 translocations: nonidentical break-points on chromosomes 11 and 22.
Proc Natl Acad Sci U S A **83**: 6122-6126
- Grossman MH, Emanuel BS, Budarf ML**
1992 Chromosomal mapping of the human catechol-O-methyltransferase gene to 22q11.1-q11.2.
Genomics **12**: 822-825
- Guarguaglini G, Battistoni A, Pittoggi C, Di Matteo G, Di Fiore B, Lavia P**
1997 Expression of the murine RanBP1 and H89-c genes is regulated from a shared bidirectional promoter during cell cycle progression.
Biochem J **325**: 277-286
- Gumpel-Pinot**
1984 Muscle and skeleton of limbs and body wall. In: NM LeDouarin, A McClaren (Eds).
Chimeras in Developmental Biology. Academic Press, New York: 281-310
- Guthrie S, Lumsden A**
1991 Formation and regeneration of rhombomere boundaries in the developing chick hindbrain.
Development **112**: 221-229
- Guthrie S, Butcher M, Lumsden A**
1991 Patterns of cell division and interkinetic nuclear migration in the chick embryo hindbrain.
J Neurobiol **22**: 742-754
- Guthrie S, Prince V, Lumsden A**
1993 Selective dispersal of avian rhombomere cells in orthotopic and heterotopic grafts.
Development **118**: 527-538
- Halford S, Lindsay E, Nayudu M, Carey AH, Baldini A, Scambler PJ**
1993a Low-copy-number repeat sequences flank the D:George/Velo-cardio-facial syndrome loci at 22q11.
Hum Mol Genet **2**: 191-196
- Halford S, Wade R, Roberts C, Daw SCM, Whiting JA, O'Donnell H, Dunham I, Bentley D, Lindsay E, Baldini, Francis F, Lehrach H, Williamson R, Wilson DJ, Goodship J, Cross I, Burn J, Scambler PJ**
1993b Isolation of a putative transcriptional regulator from the region of 22q11 deleted in D:George syndrome, Shprintzen syndrome and familial congenital heart disease.
Hum Mol Genet **2**: 2099-2107
- Halford S, Wilson DJ, Daw SCM, Roberts C, Wade R, Kamath S, Wickremasinghe A, Burn J, Goodship J, Mattel M-G, Moerman AFM, Scambler PJ**
1993c Isolation of a gene expressed during early embryogenesis from the region of 22q11 commonly deleted in D:George syndrome.
Hum Mol Genet **2**: 1577-1582
- Hall BK**
1978a Initiation of osteogenesis by mandibular mesenchyme.
Arch Oral Biol **23**: 1157-1161

- Keynes RJ, Stern CD**
1984 Segmentation in the vertebrate nervous system.
Nature **310**: 786-789
- Keynes RJ, Stern CD**
1988 Mechanisms of vertebrate segmentation.
Development **103**: 413-429
- Khoury MJ, Moore CA, Evans JA**
1994 On the use of the term "syndrome" in clinical genetics and birth defects epidemiology.
Am J Med Genet **49**: 26-28
- Kil SH, Lallier T, Bronner-Fraser M**
1996 Inhibition of cranial neural crest adhesion in vitro and migration in vivo using integrin antisense oligonucleotides.
Dev Biol **179**: 91-101
- Kimmelman CP, Denneny JC**
1982 Optic (G) syndrome.
Int J Pediatr Otorhinolaryngol **4**: 343-347
- Kinouchi A, Mori K, Ando M, Takao A**
1976 Facial appearance of patients with conotruncal abnormalities.
Pediatr Jpn **17**: 84
- Kirby ML, Stewart DE**
1983 Neural crest origin of cardiac ganglion cells in the chick embryo: identification and extirpation.
Dev Biol **97**: 433-443
- Kirby ML, Waldo KL**
1990 Role of neural crest in congenital heart disease.
Circulation **82**: 332-340
- Kirby ML, Waldo KL**
1995 Neural crest and cardiovascular patterning.
Circ Res **77**: 211-215
- Kirby ML, Gale TF, Stewart DE**
1983 Neural crest cells contribute to normal aorticopulmonary septation.
Science **220**: 1059-1061
- Kirby ML, Turnage KL 3rd, Hays BM**
1985 Characterization of conotruncal malformations following ablation of "cardiac" neural crest.
Anat Rec **213**: 87-93
- Kirby ML, Cheng G, Stadt H, Hunter G**
1995 Differential expression of the L10 ribosomal protein during heart development.
Biochem Biophys Res Commun **212**: 461-465
- Kirby ML, Hunt P, Wallis K, Thorogood P**
1997 Abnormal patterning of the aortic arch arteries does not evoke cardiac malformations.
Dev Dyn **208**: 34-47
- Kirkpatrick JA Jr, DiGeorge AM**
1968 Congenital absence of the thymus.
Am J Roentgenol Radium Ther Nucl Med **103**: 32-37
- Kirov N, Shtilbans A, Rushlow C**
1998 Isolation and characterization of a new gene encoding a member of the HIRA family of proteins from *Drosophila melanogaster*.
Gene **212**: 323-332
- Kishino T, Lalonde M, Wagstaff J**
1997 UBE3A/E6-AP mutations cause Angelman syndrome.
Nat Genet **15**: 70-73
- Koehler U, Bigoni F, Wienberg J, Stanyon R**
1995a Genomic reorganization in the concolor gibbon (*Hylobates concolor*) revealed by chromosome painting.
Genomics **30**: 287-292
- Koehler U, Arnold N, Wienberg J, Tofanelli S, Stanyon R**
1995b Genomic reorganization and disrupted chromosomal synteny in the samang (*Hylobates syndactylus*) revealed by fluorescence in situ hybridization.
Am J Phys Anthropol **97**: 37-47
- Komachi K, Redd MJ, Johnson AD**
1994 The WD repeats of Tup1 interact with the homeo domain protein 2.
Genes Dev **8**: 2857-2867
- Kontges G, Lumsden A**
1996 Rhombencephalic neural crest segmentation is preserved throughout craniofacial ontogeny.
Development **122**: 3229-3242
- Krantz ID, Piccoli DA, Spinner NB**
1997a Alagille syndrome.
J Med Genet **34**: 152-157
- Krantz ID, Rand EB, Genin A, Hunt P, Jones M, Louis AA, Graham JM Jr, Bhatt S, Piccoli DA, Spinner NB**
1997b Deletions of 20p12 in Alagille syndrome: frequency and molecular characterization.
Am J Med Genet **70**: 80-86
- Kretschmer R, Say B, Brown D, Rosen FS**
1968 Congenital aplasia of the thymus gland (DiGeorge's syndrome).
N Engl J Med **279**: 1295-1301
- Krotoski D, Domingo C, Bronner-Fraser M**
1986 Distribution of a putative cell surface receptor for fibronectin and laminin in the avian embryo.
J Cell Biol **103**: 1061-1072
- Krull CE, Collazo A, Fraser SE, Bronner-Fraser M**
1995 Segmental migration of trunk neural crest: time-lapse analysis reveals a role for PNA-binding molecules.
Development **121**: 3733-3743
- Krull CE, Lansford R, Gale NW, Collazo A, Marcelle C, Yancopoulos GD, Fraser SE, Bronner-Fraser M**
1997 Interactions of Eph-related receptors and ligands confer rostrocaudal pattern to trunk neural crest migration.
Curr Biol **7**: 571-580
- Kruse J, Mailhammer R, Wernecke H, Falssner A, Sommer I, Goridis C, Schachner M**
1984 Neural cell adhesion molecules and myelin-associated glycoprotein share a common carbohydrate moiety recognized by monoclonal antibodies L2 and HNK-1.
Nature **311**: 153-155
- Kurahashi H, Akagi K, Karakawa K, Nakamura T, Dumanski JP, Sano T, Okada S, Takai SI, Nishisho I**
1994 Isolation and mapping of cosmid markers on human chromosome 22, including one within the submicroscopically deleted region of DiGeorge syndrome.
Hum Genet **93**: 248-254
- Kurahashi H, Akagi K, Inazawa J, Ohta T, Niikawa N, Kayatani F, Sano T, Okada S, Nishisho I**
1995 Isolation and characterization of a novel gene deleted in DiGeorge syndrome.
Hum Mol Genet **4**: 541-549
- Kurahashi H, Nakayama T, Osugi Y, Tsuda E, Masuno M, Imaizumi K, Kamiya T, Sano T, Okada S, Nishisho I**
1996 Deletion mapping of 22q11 in CA1CH22 syndrome: identification of a second critical region.
Am J Hum Genet **58**: 1377-1381
- Kurahashi H, Tsuda E, Kohama R, Nakayama T, Masuno M, Imaizumi K, Kamiya T, Sano T, Okada S, Nishisho I**
1997 Another critical region for deletion of 22q11: a study of 100 patients.
Am J Med Genet **72**: 180-185
- Kuratani S, Bockman DE**
1990 The participation of neural crest derived mesenchymal cells in development of the epithelial primordium of the thymus.
Arch Histol Cytol **53**: 267-273
- Kuratani SC, Kirby ML**
1991 Initial migration and distribution of the cardiac neural crest in the avian embryo: an introduction to the concept of the circumpharyngeal crest.
Am J Anat **191**: 215-227

- Kuratani S, Martin JF, Wawersik S, Lilly B, Eichele G, Olson EN**
1994 The expression pattern of the chick homeobox gene *gHox* suggests a role in patterning of the limbs and face and in compartmentalization of somites.
Dev Biol **161**: 357-369
- Kuratani S, Matsuo I, Aizawa S**
1997 Developmental patterning and evolution of the mammalian viscerocranium: genetic insights into comparative morphology.
Dev Dyn **209**: 139-155
- Kurihara Y, Kurihara H, Suzuki H, Kodama T, Maemura K, Nagai R, Oda H, Kuwaki T, Cao YH, Kamada N, et al**
1994 Elevated blood pressure and craniofacial abnormalities in mice deficient in endothelin-1.
Nature **368**: 703-710
- Kurihara Y, Kurihara H, Maemura K, Kuwaki T, Kumada M, Yazaki Y**
1995a Impaired development of the thyroid and thymus in endothelin-1 knockout mice.
J Cardiovasc Pharmacol **26 Suppl 3**: S13-S16
- Kurihara Y, Kurihara H, Oda H, Maemura K, Nagai R, Ishikawa T, Yazaki Y**
1995b Aortic arch malformations and ventricular septal defect in mice deficient in endothelin-1.
J Clin Invest **96**: 293-300
- Kusafuka T, Wang Y, Puri P**
1996 Novel mutations of the endothelin-B receptor gene in isolated patients with Hirschsprung's disease.
Hum Mol Genet **5**: 347-349
- Kusafuka T, Wang Y, Puri P**
1997 Mutation analysis of the RET, the endothelin-B receptor, and the endothelin-3 genes in sporadic cases of Hirschsprung's disease.
J Pediatr Surg **32**: 501-504
- Lachman HM, Morrow B, Shprintzen R, Velt S, Parsia SS, Faedda G, Goldberg R, Kucherlapati R, Papalos DF**
1996 Association of codon 108/158 catechol-O-methyltransferase gene polymorphism with the psychiatric manifestations of velo-cardio-facial syndrome.
Am J Med Genet **67**: 468-472
- Lahav R, Ziller C, Dupin E, Le Douarin NM**
1996 Endothelin 3 promotes neural crest cell proliferation and mediates a vast increase in melanocyte number in culture.
Proc Natl Acad Sci U S A **93**: 3892-3897
- Lai MMR, Scriven PN, Ball C, Berry AC**
1992 Simultaneous partial monosomy 10p and trisomy 5q in a case of hypoparathyroidism.
J Med Genet **29**: 586-588
- Lallier T, Deutzmann R, Perris R, Bronner-Fraser M**
1994 Neural crest cell interactions with laminin: structural requirements and localization of the binding site for alpha 1 beta 1 integrin.
Dev Biol **162**: 451-464
- Lammer EJ, Chen DT, Hoar RM, Agnish ND, Benke PJ, Braun JT, Curry CJ, Fernhoff PM, Grix A.W Jr, Lott IT, Richard JM, Sun SC**
1985 Retinoic acid embryopathy.
New Eng J Med **313**: 837-84.
- Lammer EJ, Opitz JM**
1986 The DiGeorge anomaly is a developmental field defect.
Am J Med Genet (Suppl) **2**: 113-127
- Lamour Y, Lécluse Y, Desmaze C, Spector M, Bodescot M, Aurias A, Osley MA, Lipinski M**
1995 A human homolog of the *S. cerevisiae* HIR1 and HIR2 transcriptional repressors cloned from the DiGeorge syndrome critical region.
Hum Mol Gen **4**: 791-799
- Landolt RM, Vaughan L, Winterhalter KH, Zimmermann DR**
1995 Versican is selectively expressed in embryonic tissues that act as barriers to neural crest cell migration and axon outgrowth.
Development **121**: 2303-2312
- LANOUÉ L, Dehart DB, Hinsdale ME, Maeda N, Tint GS, Sulik KK**
1997 Limb, genital, CNS, and facial malformations result from gene/environment-induced cholesterol deficiency: further evidence for a link to sonic hedgehog.
Am J Med Genet **73**: 24-31
- Larsen WJ (Ed)**
1975
Human Embryology. Churchill Livingstone, New York
- LAYER PG, AlBER R**
1990 Patterning of chick brain vesicles as revealed by peanut agglutinin and cholinesterases.
Development **109**: 613-624
- Le Douarin NM, Jotereau**
1975 Tracing of cells of the avian thymus through embryonic life in interspecific chimeras.
J Exp Med **142**: 17-40
- Le Douarin NM, Teitel M**
1973 The migration of neural crest cells to the wall of the digestive tract in avian embryos.
J Embryol Exp Morphol **30**: 31-48
- Le Douarin NM, Dupin E, Ziller C**
1994 Genetic and epigenetic control in neural crest development.
Curr Opin Genet Dev **4**: 685-695
- Le Douarin NM**
1983
The neural crest. Cambridge University Press, Cambridge
- Le Lieve CS, Le Douarin NM**
1975 Mesenchymal derivatives of the neural crest: analysis of chimaeric quail and chick embryos.
J Embryol exp Morph **34**: 125-154
- Leblanc GG, Holbert TE, Darland T**
1995 Role of the transforming growth factor-beta family in the expression of cranial neural crest-specific phenotypes.
J Neurobiol **26**: 497-510
- Lee KF, Simon H, Chen H, Bates B, Hung MC, Hauser C**
1996 Requirement for neuregulin receptor erbB2 in neural and cardiac development.
Nature **378**: 394-398
- Lee YM, Osumi-Yamashita N, Ninomiya Y, Moon CK, Eriksson U, Eto K**
1995 Retinoic acid stage-dependently alters the migration pattern and identity of hindbrain neural crest cells.
Development **121**: 825-837
- Lekanne Deprez RH, van Biezen NA, Heutink P, Boeijharat KRGS, de Klein A, Geurts van Kessel AHM, Zwarthoff EC**
1991 A new polymorphic probe on chromosome 22; NB84 (D22S183).
Nucleic Acids Res **19**: 687
- Lengele B, Schowing J, Dhem A**
1996 Embryonic origin and fate of chondroid tissue and secondary cartilages in the avian skull.
Anat Rec **246**: 377-393
- Levy A, Demczuk S, Aurias A, Depetris D, Mattel MG and Philip N**
1995 Interstitial 22q11.1 microdeletion excluding the ADU breakpoint in a patient with the DiGeorge syndrome.
Hum Mol Genet **4**: 2417-2419
- Lewis E**
1978 A gene complex controlling segmentation in *Drosophila*.
Nature **27**: 565-570

- Li QY, Newbury-Ecob RA, Terret JA, Wilson DA, Curtis ARJ, Yi CH, Gebuhr T, Bullen PJ, Robson SC, Strachan T, Bonnet D, Lyonnet S, Young ID, Raeburn JA, Buckler AJ, Law DJ, Brook JD**
1997 Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family.
Nat Genet **15**: 21-29
- Liem KF Jr, Tremml G, Roelink H, Jessell TM**
1995 Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm.
Cell **82**: 969-979
- Lindblom A, Sandelin K, Iselius L, Dumanski J, White I, Nordenskjold M, Larsson C**
1994 Predisposition for breast cancer in carriers of constitutional translocation 11q;22q.
Am J Hum Genet **54**: 871-876
- Lindsay EA, Halford S, Wade R, Scambler PJ, Baldini A**
1993 Molecular cytogenetic characterization of the DiGeorge syndrome region using fluorescence in situ hybridization.
Genomics **17**: 403-407
- Lindsay EA, Goldberg R, Jurecik V, Morrow B, Carlson C, Kucherlapati RS, Shprintzen RJ, Baldini A**
1995 Velo-cardio-facial syndrome: frequency and extent of 22q11 deletions.
Am J Med Genet **57**: 514-522
- Lindsay EA, Rizzu P, Antonacci R, Juecic V, Delmas-Mata J, Lee CC, Kim UJ, Scambler PS, Baldini A**
1996 A transcription map in the CATCH22 critical region: identification, mapping, and ordering of four novel transcripts expressed in heart.
Genomics **32**: 104-112
- Lindsay EA, Harvey EL, Scambler PJ, Baldini A**
1998 ESZ, a gene deleted in DiGeorge syndrome, encodes a nuclear protein and is expressed during early mouse development, where it shares an expression domain with a Goosecoid-like gene.
Hum Mol Genet **7**: 629-635
- Lindsay EA, Botta A, Jurecik V, Carattini-Rivera S, Cheah YC, Rosenblatt HM, Bradley A, Baldini A**
1999 Congenital disease in mice deficient for the DiGeorge syndrome region.
Nature **401**: 379-383
- Lipson A, Fagan K, Colley A, Colley P, Sholler G, Issacs D, Oates RK**
1996 Velo-cardio-facial and partial DiGeorge phenotype in a child with interstitial deletion at 10p13—implications for cytogenetics and molecular biology.
Am J Med Genet **65**: 304-308
- Lischner HW**
1972 DiGeorge syndrome(s).
J Pediatr **81**: 1042-1044
- Llavadot R, Estivill X, Scambler P, Pritchard M**
1998a Isolation and genomic characterization of the TUPLE1/HIRA gene of the pufferfish *Fugu rubripes*.
Gene **208**: 279-283
- Llavadot R, Marques G, Pritchard M, Estivill X, Ferrus A, Scambler P**
1998b Cloning, chromosome mapping and expression analysis of the HIRA gene from *Drosophila melanogaster*.
Biochem Biophys Res Commun **249**: 486-491
- Lo L, Guillemot F, Joyner AL, Anderson DJ**
1994 MASH-1: a marker and a mutation for mammalian neural crest development.
Perspect Dev Neurobiol **2**: 191-201
- Long KR, Trofatter JA, Ramesh V, McCormick MK, Buckler AJ**
1996 Cloning and characterization of a novel human cation heavy chain gene (CLTCL).
Genomics **35**: 466-472
- Lopez JA, Chung DW, Fujikawa K, Hagen FS, Davie EW, Roth GJ**
1988 The alpha and beta chains of human platelet glycoprotein Ib are both transmembrane proteins containing a leucine-rich amino acid sequence.
Proc Natl Acad Sci USA **85**: 2135-2139
- Lorain S, Demczuk S, Lamour V, Toth S, Aurias A, Roe BA, Lipinski M**
1996 Structural organization of the WD40 repeat protein-encoding gene HIRA in the DiGeorge syndrome critical region of human chromosome 22.
Genome Res **6**: 43-50
- Lorain S, Quivy JP, Monier-Gavelle F, Scamps C, Lecluse Y, Almouzni G, Lipinski M**
1998 Core histones and HIRIP3, a novel histone-binding protein, directly interact with WD repeat protein HIRA.
Mol Cell Biol **18**: 5546-5556
- Lu-Kuo J, Ward DC, Spritz RA**
1993 Fluorescence in situ hybridization mapping of 25 markers on distal human chromosome 2q surrounding the human Waardenburg syndrome, type 1 (WS1) locus (PAX3 gene).
Genomics **16**: 173-179
- Ludecke HJ, Wagner MJ, Hardmann J, La Pillo B, Parrish JE, Willems PJ, Haan EA, Frydman M, Hamers GJ, Wells DE, et al**
1995 Molecular dissection of a contiguous gene syndrome: localization of the genes involved in the Langer-Giedion syndrome.
Hum Mol Genet **4**: 31-36
- Lumsden A**
1991 Cell lineage restrictions in the chick embryo hindbrain.
Philos Trans R Soc Lond B Biol Sci **331**: 281-6
- Lumsden A, Guthrie S**
1991 Alternating patterns of cell surface properties and neural crest cell migration during segmentation of the chick hindbrain.
Development Suppl **2**: 9-15
- Lumsden A, Keynes R**
1989 Segmental patterns of neuronal development in the chick hindbrain.
Nature **337**: 424-428
- Lumsden A, Sprawson N, Graham A**
1991 Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo.
Development **113**: 1281-1291
- Lund J, Roe B, Chen F, Budarf M, Galili N, Riblet R, Miller RD, Emanuel BS, Reeves RH**
1999 Sequence-ready physical maps of the mouse Chromosome 16 region with conserved synteny to the human Velocardiofacial syndrome region on 22q11.
Manus Genome **10**: 438-443
- Luo Y, Ceccherini I, Pasini B, Matera I, Bicocchi MB, Barone V, Boccardi R, Kaariainen H, Weber D, Devoto M, et al**
1993 Close linkage with the RET protooncogene and boundaries of deletion mutations in autosomal dominant Hirschsprung disease.
Hum Mol Genet **2**: 1803-1808
- Macdonald R, Barth KA, Xu Q, Holder N, Mikkola I, Wilson SW**
1995 Midline signaling is required for Pax gene regulation and patterning of the eyes.
Development **121**: 3267-3278
- Mager AM, Grapin-Botton A, Ladjali K, Meyer K, Wolff CM, Stiegler P, Bonnini MA, Remy P**
1998 The avian *Fl* gene is specifically expressed during embryogenesis in a subset of neural crest cells giving rise to the mesenchyme.
Int J Dev Biol **42**: 561-572
- Magnaghi P, Roberts C, Lorain S, Lipinski M, Scambler PJ**
1998 HIRA, a mammalian homologue of *Saccharomyces cerevisiae* transcriptional co-repressors, interacts with Pax3.
Nat Genet **20**: 74-77
- Magyar JP, Martini R, Ruelicke T, Aguzzi A, Adlkofer K, Dembic Z, Zielasek J, Toyka KV, Suter U**
1996 Impaired differentiation of Schwann cells in transgenic mice with increased PMP22 gene dosage.
J Neurosci **16**: 5351-5360
- Manley NR, Capecchi MR**
1995 The role of Hoxa-3 in mouse thymus and thyroid development.
Development **121**: 1989-2003
- Mansouri A, Stoykova A, Torres M, Gruss P**
1996 Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice.
Development **122**: 831-838



- Monaco G, Pignata C, Rossi E, Mascellaro O, Coccoza S, Cicciarra F**
1991 DiGeorge anomaly associated with 10p deletion.
Am J Med Genet **39**: 215-216
- Monier-Gavelle F, Duband JL**
1995 Control of N-cadherin mediated intercellular adhesion in migrating neural crest cells in vitro.
J Cell Sci **108**: 3839-3853
- Monier-Gavelle F, Duband JL**
1997 Cross talk between adhesion molecules: control of N-cadherin activity by intracellular signals elicited by beta1 and beta3 integrins in migrating neural crest cells.
Cell Biol **137**: 1663-1681
- Montini E, Rugarli E, Van de Yosse E, Andolfi G, Mariani M, Puca AA, Consalez GG, den Dunnen JT, Ballabio A, Franco B**
1997 A novel human serine-threonine phosphatase related to the Drosophila retinal degeneration C (rdgC) gene is selectively expressed in sensory neurons of neural crest origin.
Hum Mol Genet **6**: 1137-1145
- Moore MA, Owen JJ**
1967 Experimental studies on the development of the thymus.
J Exp Med **126**: 715-726
- Moore MW, Klein RD, Farinas I, Sauer H, Armanini M, Phillips H, Reichardt LF, Ryan AM, Carver-Moore K, Rosenthal A**
1996 Renal and neuronal abnormalities in mice lacking GDNF.
Nature **382**: 76-79
- Morescalchi MA, Schempp W, Consigliere S, Bigoni F, Wienberg J, Stanyon R**
1997 Mapping chromosomal homology between humans and the black-handed spider monkey by fluorescence in situ hybridization.
Chromosome Res **5**: 527-536
- Mori T, Miura K, Fujiwara T, Shin S, Inazawa J, and Nakamura Y**
1996 Isolation and mapping of a human gene (DIFF6) homologous to yeast CDC3, CDC10, CDC11 and CDC12, and mouse D:6.
Cytogenet Cell Genet **73**: 224-227
- Morriss-Kay G**
1993 Retinoic acid and craniofacial development: molecules and morphogenesis.
Bioessays **15**: 9-15
- Morriss-Kay GM**
1996 Craniofacial defects in AP-2 null mutant mice.
Bioessays **18**: 785-788
- Morriss-Kay GM, Sokolova H**
1996 Embryonic development and pattern formation.
FASEB J **10**: 961-968
- Morrow B, Goldberg R, Carlson C, Gupta RD, Sirotkin H, Collins J, Dunham I, O'Donnell H, Scambler P, Shprintzen R, Kucherlapati R**
1995 Molecular definition of the 22q11 deletions in velo-cardio-facial syndrome.
Am J Hum Genet **56**: 1391-1403
- Moury JD, Jacobson AG**
1989 Neural fold formation at newly created boundaries between neural plate and epidermis in the axolotl.
Dev Biol **133**: 44-57
- Moynihan TP, Ardley HC, Leek JP, Thomson J, Brindle NS, Markham AF, Robindson PA**
1996 Characterization of a human ubiquitin-conjugating enzyme gene UBE2L3.
Mamm Genome **7**: 520-525
- Mulder MP, Wilke M, Langeveld A, Wilming LG, Hagemeljer A, van Druenen E, Zwarthoff EC, Riegman P, Beelen WH, van den Ouweland AMW, Halley DJJ, Meljers C**
1995 Positional mapping of loci in the DiGeorge critical region at chromosome 22q11 using a new marker (D22S183).
Hum Genet **96**: 133-141
- Muller S, O'Brien PC, Ferguson-Smith MA, Wienberg J**
1997 Reciprocal chromosome painting between human and prosimians (*Eulemur macaco macaco* and *E. fulvus mayottensis*).
Cytogenet Cell Genet **78**: 260-271
- Mulligan LM, Kwok JB, Healey CS, Eldson MJ, Eng C, Gardner E, Lloye DR, Mole SE, Moore JK, Papi L, et al**
1993 Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A.
Nature **363**: 458-460
- Muraoka M, Konishi M, Kikuchi-Yanoshita R, Tanaka K, Shitara N, Chong JM, Iwama T, Miyaki M**
1996 p300 gene alterations in colorectal and gastric carcinomas.
Oncogene **12**: 1565-1569
- Murdoch-Kinch CA, Ward RE**
1997 Metacarpophalangeal analysis in Crouzon syndrome: additional evidence for phenotypic convergence with the acrocephalosyndactyly syndromes.
Am J Med Genet **73**: 61-66
- Murphy M, Reid K, Ford M, Furness JB, Bartlett PF**
1994 FGF2 regulates proliferation of neural crest cells, with subsequent neuronal differentiation regulated by UF or related factors.
Development **120**: 3519-3528
- Nakagawa S, Takeichi M**
1995 Neural crest cell-cell adhesion controlled by sequential and subpopulation-specific expression of novel cadherins.
Development **121**: 1321-1332
- Nakagawa S, Takeichi M**
1998 Neural crest emigration from the neural tube depends on regulated cadherin expression.
Development **125**: 2963-2971
- Nakata K, Nagai T, Aruga J, Mikoshiba K**
1997 *Xenopus zic1*, a primary regulator both in neural and neural crest development.
Proc Natl Acad Sci U S A **94**: 11980-11985
- Nataf Y, Lecoin L, Eichmann A, Le Douarin NM**
1996 Endothelin-B receptor is expressed by neural crest cells in the avian embryo.
Proc Natl Acad Sci U S A **93**: 9645-9650
- Neer EJ, Schmidt CJ, Nambudripad R, Smith TF**
1994 The ancient regulatory-protein family of WD-repeat proteins.
Nature **371**: 297-300
- Neill CA**
1987 Congenital cardiac malformations and syndromes. In: Pierpont ME, Mofer JH (Eds).
Genetics of cardiovascular disease. Martinus Nijhoff, Boston: 95-112
- Neuhauss SC, Solnica-Krezel L, Schier AF, Zwartkruis F, Stemple DL, Malicki J, Abdellah S, Stainier DY, Driever W**
1996 Mutations affecting craniofacial development in zebrafish.
Development **123**: 357-367
- Newgreen DF, Minichiello J**
1995 Control of epithel-mesenchymal transformation. I. Events in the onset of neural crest cell migration are separable and inducible by protein kinase inhibitors.
Dev Biol **170**: 91-101
- Newman CS, Grow MW, Cleaver O, Chia F, Krieg P**
1997 Xbp, a vertebrate gene related to bagpipe, is expressed in developing craniofacial structures and in anterior gut muscle.
Dev Biol **181**: 223-233
- Nichols DH**
1981 Neural crest cell formation in the head of the mouse embryo as observed using a new histological technique.
J Embryol Exp Morphol **64**: 105-120

- Perris R, Brandenberger R, Chiquet M**
1996 Differential neural crest cell attachment and migration on avian laminin isoforms.
Int J Dev Neurosci **14**: 297-314
- Petrij F, Giles RH, Dauwerse HG, Saris JJ, Hennekam RC, Masuno M, Tommerup N, van Ommen GJ, Goodman RH, Peters DJ, et al**
1995 Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP.
Nature **376**: 348-351
- Pichel JG, Shen L, Sheng HZ, Granholm AC, Drago J, Grinberg A, Lee EJ, Huang SP, Saarma M, Hoffer BJ, Sariola H, Westphal H**
1996 Defects in enteric innervation and kidney development in mice lacking GDNF.
Nature **382**: 73-76
- Pignata C, D'Agostino A, Finelli P, Fiore M, Scotese I, Cosentini E, Cuomo C, Venuta S**
1996 Progressive deficiencies in blood T cells associated with a 10p12-13 interstitial deletion.
Clin Immunol Immunopathol **80**: 9-15
- Piotrowski T, Schilling TF, Brand M, Jlang YJ, Heisenberg CP, Beuchle D, Grandel H, van Eeden FJ, Furutani-Seiki M, Granato M, Haffter P, Hammerschmidt M, Kane DA, Kelsh RN, Mullins MC, Odenthal J, Warga RM, Nusslein-Volhard C**
1996 Jaw and branchial arch mutants in zebrafish II: anterior arches and cartilage differentiation.
Development **123**: 345-356
- Pizzuti A, Novelli G, Mari A, Ratti A, Colosimo A, Amati F, Penso D, Sanguuolo F, Calabrese G, Palka G, Salani V, Gennarelli M, Mingarelli R, Scarlato G, Scambler P, Dallapiccola B**
1996 Human homologue sequences to the *Drosophila* dishevelled segment polarity gene are deleted in the DiGeorge syndrome.
Am J Hum Genet **58**: 722-729
- Pizzuti A, Novelli G, Ratti A, Amati F, Aari M, Calabrese G, Nicolis S, Silani V, Marino B, Scarlato G, Ottolenghi S, Dallapiccola B**
1997 UFDIL, a developmentally expressed ubiquitination gene, is deleted in CATCH 22 syndrome.
Hum Mol Genet **6**: 259-265
- Poelmann RE, Mentink MM, Gittenberger-de Groot AC**
1994 Rostral-caudal polarity in the avian somite related to paraxial segmentation. A study on HNK-1, tenascin and neurofilament expression.
Anat Embryol (Berl) **190**: 101-111
- Poelmann RE, Mikawa T, Gittenberger-de Groot**
1998 Neural crest cells in outflow tract septation of the embryonic chicken heart: differentiation and apoptosis.
Dev Dyn **212**: 373-384
- Poole TJ, Thierry JP**
1986 Anti-bodies and synthetic peptides that block cell-fibronectin adhesion arrest neural crest migration in vivo.
Prog Clin Biol Res **217B**: 235-238
- Prasad C, Quackenbush EJ, Whiteman D, Korf B**
1997 Limb anomalies in DiGeorge and CHARGE syndromes.
Am J Med Genet **68**: 179-181
- Prince V, Lumsden A**
1994 Hox-2 expression in normal and transposed rhombomeres: independent regulation in the neural tube and neural crest.
Development **120**: 911-923
- Promisel Cooper J, Roth S, Simpson RT**
1994 The global transcriptional regulators SSN6 and TUP1, play distinct roles in the establishment of a repressive chromatin structure.
Genes & Dev **8**: 1400-1410
- Puech A, Saint-Jore B, Funke B, Gilbert DJ, Sirotkin H, Copeland NG, Jenkins NA, Kucherlapati R, Morrow B, Skoultschi AI**
1997 Comparative mapping of the human 22q11 chromosomal region and the orthologous region in mice reveals complex changes in gene organization.
Proc Natl Acad Sci USA **94**: 14608-14613
- Quaderi NA, Schwelger S, Gaudenz K, Franco B, Rugarli EI, Berger W, Feldman GJ, Volta M, Andolfi G, Gilgenkrantz S, Marion RW, Hennekam RC, Opitz JM, Muenke M, Ropers HH, Ballabio A**
1997 Opitz G/BBB syndrome, a defect of midline development, is due to mutations in a new RING finger gene on Xp22.
Nat Genet **17**: 285-291
- Qiu M, Bulfone A, Martinez S, Meneses JJ, Shimamura K, Pedersen RA, Rubenstein JL**
1995 Null mutation of *Dlx-2* results in abnormal morphogenesis of proximal first and second branchial arch derivatives and abnormal differentiation in the forebrain.
Genes Dev **9**: 2523-2538
- Rauch A, Hofbeck M, Lelpold G, Klinge J, Trautmann U, Kirsch M, Singer H, Pfeiffer RA**
1998 Incidence and significance of 22q11 hemizygosity in patients with interrupted aortic arch.
Am J Med Genet **78**: 322-331
- Raudsepp T, Fronicke L, Scherthan H, Gustavsson I, Chowdhary BP**
1996 Zoo-FISH delineates conserved chromosomal segments in horse and man.
Chromos Res **4**: 218-225
- Raven CP**
1931 Die eigentümliche Bildungsweise des Neuralrohres beim Axotoll und die Lage des Ganglienleitenmaterials.
Anat Anz **71**: 161-166
- Read AP, Newton VE**
1997 Waardenburg syndrome.
J Med Genet **34**: 656-665
- Reaume AG, de Sousa PA, Kulkarni S, Langille BL, Zhu D, Davies TC, Juneja SC, Kidder GM, Rossant J**
1995 Cardiac malformation in neonatal mice lacking connexin43.
Science **267**: 1831-1834
- Reedy MV, Faraco CD, Erickson CA**
1998 The delayed entry of thoracic neural crest cells into the dorsolateral path is a consequence of the late emigration of melanogenic neural crest cells from the neural tube.
Dev Biol **200**: 234-246
- Rettenberger G, Klett C, Zechner U, Bruch J, Just W, Vogel W, Hameister H**
1995a Zoo-FISH analysis: cat and human karyotypes closely resemble the putative ancestral mammalian karyotype.
Chromosome Res **3**: 479-486
- Rettenberger G, Klett C, Zechner U, Kunz J, Vogel W, Hameister H**
1995b Visualization of the conservation of synteny between humans and pigs by heterologous chromosomal painting.
Genomics **26**: 372-378
- Reynolds AB, Herbert L, Cleveland JL, Berg ST, Saut JR**
1992 p120, a novel substrate of protein tyrosine kinases and of p60^{src}, is related to cadherin-binding factors beta-catenin, plakoglobin and armadillo.
Oncogene **7**: 2439-2445
- Reynolds AB, Daniel J, McCrean PD, Wheelock MJ, Wu J, Zhang Z**
1994 Identification of a new catenin: the tyrosine kinase substrate p120cas associates with E-cadherin complexes.
Mol Cell Biol **14**: 8333-8342
- Reynolds AB, Jenkins NA, Gilbert DJ, Copeland NG, Shapiro DN, Wu J, Daniel JM**
1996 The gene encoding p120cas, a novel catenin, localizes on human chromosome 11q11 (CTNND) and mouse chromosome 2 (Cttns).
Genomics **31**: 127-129
- Richard F, Lombard M, Dutrillaux B**
1996 Zoo-FISH suggests a complete homology between human and capuchin monkey (*Platyrrhini*) euchromatin.
Genomics **36**: 417-423
- Richardson MK, Hanken J, Gooneratne ML, Pieau C, Raynaud A, Aelwood L, Wright GM**
1997 There is no highly conserved stage in embryonic vertebrates: implications for current theories of evolution and development.
Anat Embryol (Berl) **196**: 91-106

- Richman JM**
1992 The role of retinoids in normal and abnormal embryonic craniofacial morphogenesis.
Crit Rev Oral Biol Med 4: 93-109
- Rickman M, Fawcett J, Keynes RJ**
1985 The migration of neural crest cells and the growth of motor axons through the rostral half of the chick somite.
J Embryol exp Morph 90: 437-455
- Rijli FM, Mark M, Lakkaraju S, Dierich A, Dolle P, Chambon P**
1993 A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene.
Cell 75: 1333-1349
- Rijli FM, Chambon P**
1997 Genetic interactions of Hox genes in limb development: learning from compound mutants.
Curr Opin Genet Dev 7: 481-487
- Rijli FM, Gavalas A, Chambon P**
1998 Segmentation and specification in the branchial region of the head: the role of the Hox selector genes.
Int J Dev Biol 42: 393-401
- Ring C, Hassell J, Halfter W**
1996 Expression pattern of collagen IX and potential role in the segmentation of the peripheral nervous system.
Dev Biol 180: 41-53
- Rivera-Perez JA, Mallo M, Gendron-Maguire M, Gridley T, Behringer RR**
1995 Goosecoid is not an essential component of the mouse gastrula organizer but is required for craniofacial and rib development.
Development 121: 3005-312
- Rizzu P, Lindsay EA, Taylor C, O'Donnell H, Levy A, Scambler P, Baldini A**
1996 Cloning and comparative mapping of a gene from the commonly deleted region of DiGeorge and Velocardiofacial syndromes conserved in *C. Elegans*, *Manim*
Genome 7: 639-643
- Roberts C, Daw SC, Halford S, Scambler PJ**
1997 Cloning and developmental expression analysis of chick Hira (Chira), a candidate gene for DiGeorge syndrome.
Hum Mol Genet 6: 237-245
- Robertson K, Mason I**
1995 Expression of ret in the chicken embryo suggests roles in regionalisation of the vagal neural tube and somites and in development of multiple neural crest and placodal fociages.
Mech Dev 53: 329-344
- Robertson K, Mason I, Hall S**
1997 Hirschsprung's disease: genetic mutations in mice and men.
Gut 41: 436-441
- Robin NH, Feldman GJ, Aronson AL, Mitchell HF, Weksberg R, Leonard CO, Burton BK, Josephson KD, Laxova R, Aleck KA, Alfanson JE, Guinalmeida ML, Martin RA, Leichtman LG, Price RA, Opitz JM, Muenke M**
1995 Optz syndrome is genetically heterogenous with one locus on Xp22, and a second locus on 22q11.2.
Nat Genet 11: 459-461
- Robinson GW, Mahon KA**
1994 Differential and overlapping expression domains of Dlx-2 and Dlx-3 suggest distinct roles for Distal-less homeobox genes in craniofacial development.
Mech Dev 48: 199-215
- Roessler E, Belloni E, Gaudenz K, Vargas F, Scherer SW, Tsui LC, Muenke M**
1997a Mutations in the C-terminal domain of Sonic Hedgehog cause holoprosencephaly.
Hum Mol Genet 6: 1847-1853
- Roessler E, Ward DE, Gaudenz K, Belloni E, Scherer SW, Donnai D, Siegel-Bartelt J, Tsui LC, Muenke M**
1997b Cytogenetic rearrangements involving the loss of the Sonic Hedgehog gene at 7q36 cause holoprosencephaly.
Hum Genet 100: 172-181
- Romeo G, Ronchetto P, Luo Y, Barone Y, Servi M, Ceccherini I, Pasini B, Bocciardi R, Lerone M, Kaariainen H, et al**
1994 Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease.
Nature 367: 377-378
- Rouleau GA, Seizinger BR, Wertschek W, Haines JL, Superneau DW, Martuza RL, Gusella JF**
1990 Flanking markers bracket the neurofibromatosis type 2 (NF2) gene on chromosome 22.
Am J Hum Genet 46: 323-328
- Rowley JD**
1982 Identification of the constant chromosome regions involved in human hematologic malignant disease.
Science 216: 749-751
- Ryan AK, Goodship JA, Wilson DJ, Philip N, Levy A, Seidel H, Schuffenhauer S, Oechsler H, Belohradsky B, Prieur M, Aurias A, Raymond FL, Clayton-Smith J, Hatchwell E, McKeown C, Beemer FA, Dallapiccola B, Novelli G, Hurst JA, Ignatius J, Green AJ, Winter RM, Brueton L, Brondum-Nielsen K, Scambler PJ, et al**
1997 Spectrum of clinical features associated with interstitial chromosome 22q11 deletions: a European collaborative study.
J Med Genet 34: 798-804
- Rychter Z**
1978 Analysis of relations between aortic arches and aorticopulmonary septation.
Birth Defects Orig Artic Ser 14: 443-448
- Sabry MA, Zaki M, Abul Hassan SJ, Ramadan DG, Abdel Rasool MA, al Awadi SA, al Saleh Q**
1998 Kenny-Caffey syndrome is part of the CATCH 22 haploinsufficiency cluster.
J Med Genet 35: 31-36
- Sadler TW, Liu ET, Augustine KA**
1995 Antisense targeting of engrailed-1 causes abnormal axis formation in mouse embryos.
Teratology 51: 292-299
- Saglio G, Pane F, Martinelli G, Guerrasio A**
1997 BCR/ABL transcripts and leukemia phenotype: an unsolved puzzle.
Leuk Lymphoma 26: 281-286
- Saint-Jore B, Puech A, Heyer J, Lin Q, Raine C, Kucherlapati R, Skoultschi AI**
1998 Goosecoid-like (GscL), a candidate gene for velocardiofacial syndrome, is not essential for normal mouse development.
Hum Mol Genet 7: 1841-1849
- Sakurai T, Yanagisawa M, Takuwa Y, Miyazaki H, Kimura S, Goto K, Masaki T**
1990 Cloning of a cDNA encoding a non-isopeptide-selective subtype of the endothelin receptor.
Nature 348: 732-735
- Saldívar JR, Krull CE, Krumlauf R, Ariza-McNaughton L, Bronner-Fraser M**
1996 Rhombomere of origin determines autonomous versus environmentally regulated expression of Hoxa-3 in the avian embryo.
Development 122: 895-904
- Saldívar JR, Sechrist JW, Krull CE, Ruffins S, Bronner-Fraser M**
1997 Dorsal hindbrain ablation results in rerouting of neural crest migration and changes in gene expression, but normal thyroid development.
Development 124: 2729-2739
- Sanchez MP, Siles-Santiago I, Frisen J, He B, Lira SA, Barbacid M**
1996 Renal agenesis and the absence of enteric neurons in mice lacking GDNF.
Nature 382: 70-73
- Sanders SL, Field CM**
1994 Cell division. Septins in common?
Curr Biol 4: 907-910

- Sanford LP, Ormsby I, Gittenberger-de Groot AC, Sarloia H, Friedman R, Boivin GP, Cardell EL, Doetschman T**
1997 TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes.
Development **124**: 2659-2670
- Satokata I, Maas R**
1994 Mx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development.
Nat Genet **6**: 348-356
- Saunders JW Jr, Gasseling MT, Cairns JM**
1959 The differentiation of prospective thigh mesoderm grafted beneath the apical ectodermal ridge of the wing bud in the chick embryo.
Dev Biol **1**: 281-301
- Scambler PJ, Carey AH, Wyse RKH, Roach S, Dumanski JP, Nordenskjold M, Williamson R**
1991 Microdeletions within 22q11 associated with sporadic and familial DiGeorge syndrome.
Genomics **10**: 201-206
- Scambler PJ, Kelly D, Lindsay E, Williamson R, Goldberg R, Shprintzen R, Wilson DI, Goodship JA, Criss IE, Burn J**
1992 Velo-cardio-facial syndrome associated with chromosome 22 deletions encompassing the DiGeorge locus.
Lancet **339**: 1138-1139
- Scamps C, Lorain S, Lamour V, Lipinski M**
1996 The HIR protein family: isolation and characterization of a complete murine cDNA.
Biochim Biophys Acta **1306**: 5-8
- Schedl A, Ross A, Lee M, Engelkamp D, Rashbass P, van Heyningen V, Hastle ND**
1996 Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities.
Cell **86**: 71-82
- Schilling TF, Piotrowski T, Grandel H, Brand M, Heisenberg CP, Jiang YJ, Beuchle D, Hammerschmidt M, Kane DA, Mullins MC, van Eeden FJ, Keish RN, Furutani-Seiki M, Granato M, Halfter P, Odenthal J, Warga RM, Trowe T, Nusslein-Volhard C**
1996a Jaw and branchial arch mutants in zebrafish I: branchial arches.
Development **123**: 329-344
- Schilling TF, Walker C, Kimmel CB**
1996b The chinless mutation and neural crest cell interactions in zebrafish jaw development.
Development **122**: 1417-1426
- Schinzel A, Schmid W, Fraccaro M, Tiepolo L, Zuffardi O, Opitz JM, Lindsten J, Zetterqvist P, Enell H, Baccichetti C, Tenconi R, Pagon RA**
1981 The "cat eye syndrome": a centric small marker chromosome probably derived from a no.22 (tetrasomy 22pter to q11) associated with a characteristic phenotype. Report of 11 patients and delineation of the clinical picture.
Hum Genet **57**: 148-158
- Schmickel RD**
1986 Contiguous gene syndromes: a component of recognizable syndromes.
J Pediatr **109**: 231-241
- Schorle H, Meier P, Buchert M, Jaenisch R, Mitchell PJ**
1996 Transcription factor AP-2 essential for cranial closure and craniofacial development.
Nature **381**: 235-238
- Schuchardt A, D'Agati V, Larsson-Blomberg L, Costantini F, Pachnis V**
1994 Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor Ret.
Nature **367**: 380-383
- Schuffenhauer S, Seidel H, Oechster H, Belohradsky B, Bernsau U, Murken J, Meltinger T**
1995 DiGeorge syndrome and partial monosomy 10p: case report and review.
Ann Genet **38**: 162-167
- Schutz B, Niessing J**
1994 Cloning and structure of a chicken zinc finger cDNA: restricted expression in developing neural crest cells.
Gene **148**: 227-236
- Scott MP, Weiner A J**
1984 Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila.
Proc Natl Acad Sci USA **81**: 4115-4119
- Sechrist J, Scherson T, Bronner-Fraser M**
1994 Rhombomere rotation reveals that multiple mechanisms contribute to the segmental pattern of hindbrain neural crest migration.
Development **120**: 1777-1790
- Sechrist J, Nieto MA, Zamanian RT, Bronner-Fraser M**
1995 Regulative response of the cranial neural tube after neural fold ablation: spatiotemporal nature of neural crest regeneration and up-regulation of Slug.
Development **121**: 4103-4115
- Selleck MA, Bronner-Fraser M**
1995 Origins of the avian neural crest: the role of neural plate-epidermal interactions.
Development **121**: 525-538
- Serbedzija GN, Bronner-Fraser M, Fraser SE**
1989 A vital dye analysis of the timing and pathways of avian trunk neural crest cell migration.
Development **106**: 809-816
- Serbedzija GN, McMahon AP**
1997 Analysis of neural crest cell migration in Spotch mice using a neural crest-specific LacZ reporter.
Dev Biol **185**: 139-147
- Shalev SA, Dar H, Barel H, Borochowitz Z**
1996 Upper limb malformations in chromosome 22q11 deletions.
Am J Med Genet **62**: 302
- Shapira B, Borochowitz Z, Bar-EI H, Dar H, Etzioni A, Lorber A**
1994 Deletion of the short arm of chromosome 10 (10p13): report of a patient and review.
Am J Med Genet **52**: 34-38
- Sherwood PW, Tang SY, Osley MA**
1993 Characterization of HIR1 and HIR2, two genes required for regulation of histone gene transcription in *Saccharomyces cerevisiae*.
Mol Cell Biol **13**: 28-38
- Shimeld SM, McKay IJ, Sharpe PT**
1996 The murine homeobox gene Msx-3 shows highly restricted expression in the developing neural tube.
Mech Dev **55**: 201-210
- Shimizu T, Takao A, Ando M, Hirayama A**
1984 Conotruncal face syndrome: its heterogeneity and association with thymus involution. In: Nova JJ, Takao A (Eds) *Congenital Heart Disease: Causes and Processes*. Futura Publishing, Mount Kisco, NY: 29-41
- Shin MK, Levorse JM, Ingram RS, Tilghman SM**
1999 The temporal requirement for endothelin receptor-B signaling during neural crest development.
Nature **402**: 496-501
- Shprintzen RJ, Goldberg RB, Lewin ML, Sidoti EJ, Berkman MD, Argamaso RV, Young D**
1978 A new syndrome involving cleft palate, cardiac anomalies, typical faces, and learning disabilities: velo-cardio-facial syndrome.
Cleft Palate J **15**: 56-62
- Shprintzen RJ, Goldberg RB, Young D, Wolford L**
1981 The velo-cardio-facial syndrome: a clinical and genetic analysis.
Pediatrics **67**: 167-172
- Sieber-Blum M, Cohen AM**
1980 Clonal analysis of quail neural crest cells: they are pluripotent and differentiate in vitro in the absence of non-crest cells.
Dev Biol **80**: 96-106
- Sieber-Blum M (Ed)**
1998 *Neurotrophins and the Neural Crest*. CRC Press, New York

Wilson R, Ainscough R, Anderson K, Baynes C, Berks M, Bonfield J, Burton J, Connell M, Copsey T, Cooper J, Coulson A, Craxton M, Dear S, Du Z, Durbin R, Favello A, Fraser A, Fulton L, Gardner A, Green P, Hawkins T, Hillier L, Jier M, Johnston L, Jones M, Kershaw J, Kirsten J, Laisster N, Latreille P, Lightning J, Lloyd C, Mortimore B, O'Callaghan M, Parson J, Percy C, Rifken L, Roopra A, Saunders D, Shownkeen R, Sims M, Smaldon N, Smith A, Smith M, Sonhammer E, Staden R, Sulston J, Thierry-Mieg J, Thomas K, Vaudin M, Vaughan K, Waterston R, Watson A, Weinstock L, Wilkinson-Sproat J, Wohldman P
1994 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368**: 32-38

Winning RS, Sargent TD
1994 Pagliaccio, a member of the Eph family of receptor tyrosine kinase genes, has localized expression in a subset of neural crest and neural tissues in *Xenopus laevis* embryos. *Mech Dev* **46**: 219-229

Winter RM
1996 Analysing human developmental abnormalities. *Bioessays* **18**: 965-971

Winqvist R, Lundstrom K, Salminen M, Laatikainen M, Ulmanen I
1991 Mapping of human catechol-O-methyltransferase gene to 22q11.2 and detection of a frequent RFLP with BglI. *Cytogenet Cell Genet* **58**: 2051

Wirth J, Wagner T, Meyer J, Pfeiffer RA, Tietze HU, Schempp W, Scherer G
1996 Translocation break-points in three patients with campomelic dysplasia and autosomal sex reversal map more than 130 kb from Sox9. *Hum Genet* **97**: 186-193

Wunsch AM, Haas AL
1995 Ubiquitin-protein conjugates selectively distribute during early chicken embryogenesis. *Dev Dyn* **204**: 118-122

Xu D, Emoto N, Giaid A, Slaughter C, Kaw S, deWit D, Yanagisawa M
1994 ECE-1: a membrane-bound metalloprotease that catalyzes the proteolytic activation of big endothelin-1. *Cell* **78**: 473-485

Xu Q, Aldus G, Holder N, Wilkinson DG
1995 Expression of truncated Sek-1 receptor tyrosine kinase disrupts the segmental restriction of gene expression in the *Xenopus* and zebrafish hindbrain. *Development* **121**: 4005-4016

Yagi M, Edelhoff S, Disteche CM, Roth GJ
1994 Structural characterization and chromosomal location of the gene encoding human platelet glycoprotein Ib-beta. *J Biol Chem* **269**: 17424-17427

Yamada G, Mansouri A, Torres M, Stuart ET, Blum M, Schultz M, De Robertis EM, Gruss P
1995 Targeted mutation of the murine goosecoid gene results in craniofacial defects and neonatal death. *Development* **121**: 2917-2922

Yamada G, Ueno K, Nakamura S, Hanamura Y, Yasui K, Uemura M, Ezuru Y, Mansouri A, Blum M, Sugimura K
1997 Nasal and pharyngeal abnormalities caused by the mouse Goosecoid gene mutation. *Biochem Biophys Res Commun* **233**: 161-165

Yamagishi H, Ishii C, Maeda J, Kojima Y, Matsuoka R, Kimura M, Takao A, Momma K, Matsuo N
1998 Phenotypic discordance in monozygotic twins with 22q11.2 deletion. *Am J Med Genet* **78**: 319-21

Yamagishi H, Garg V, Matsuoka R, Thomas T, Srivastava D
1999 A molecular pathway revealing a genetic basis for human cardiac and craniofacial defects. *Science* **283**: 1158-1161

Yanagisawa H, Yanagisawa M, Kapur R, Richardson J, Williams S, Clouthier D, Wit D, Emoto N, Hammer R
1998 Dual genetic pathways of endothelin-mediated intercellular signaling revealed by targeted disruption of endothelin converting enzyme-1 gene. *Development* **125**: 825-836

Yang F, Carter NP, Shi L, Ferguson-Smith MA
1995 A comparative study of karyotypes of muritjacs by chromosome painting. *Chromosoma* **103**: 642-651

Zeidler R, Joos S, Delecluse HJ, Klobbeck G, Guillaume M, Lenoir GM, Bornkamm GW, Lipp M
1994 Break-points of Burkitt's lymphoma t(8;22) translocations map within a distance of 300 kb downstream of MYC. *Genes Chromosomes Cancer* **9**: 282-287

Zhang M, Kim HJ, Marshall H, Gendron-Maguire M, Lucas DA, Baron A, Gudas LJ, Gridley T, Krumlauf R, Grippo JF
1994 Ectopic Hoxa-1 induces rhombomere transformation in mouse hindbrain. *Development* **120**: 2431-2442

Zhang J, Hagopian-Donaldson S, Serbedzija G, Elsemore J, Plehn-Dujowich D, McMahon AP, Flavell RA, Williams T
1996 Neural tube, skeletal and body wall defects in mice lacking transcription factor AP-2. *Nature* **381**: 238-241

Zhang JM, Hoffmann R, Sieber-Blum M
1997 Mitogenic and anti-proliferative signals for neural crest cells and the neurogenic action of TGF-beta1. *Dev Dyn* **208**: 375-386

Zhu CC, Yamada G, Blum M
1997 Correlation between the loss of middle ear bones and altered goosecoid gene expression in the branchial region following retinoic acid treatment of mouse embryos in vivo. *Biochem Biophys Res Commun* **235**: 748-753

Zucman-Rossi J, Legoux P, Thomas G
1996 Identification of new members of the Gas2 and Ras families in the 22q12 chromosome region. *Genomics* **38**: 247-254

Zwilling E
1955 Ectoderm-mesoderm relationship in the development of the chick embryo 1 mb bud. *J Exp Zool* **128**: 423-441

- URL1 <http://www.ncbi.nlm.nih.gov/cgi-bin/SCIENCE96/chr722>
 URL2 <http://www.sanger.ac.uk/Software/Acedb/>
 URL3 <http://www.gene.ucl.ac.uk/nomenclature>
 URL4 <http://www.oncor.com>
 URL5 <http://locus.umdncj.edu/nigms/>
 URL6 <http://blochem.boehringer-mannheim.com>
 URL7 <http://www.neb.com>
 URL8 <http://www.biotech.pharmacia.se>
 URL9 <http://www.eurogentec.com>
 URL10 <http://www.apbiotech.com>
 URL11 <http://www.qiagen.com>
 URL12 <http://www.clontech.com>
 URL13 <http://www.beckman.com>
 URL14 <http://www.stratagene.com>
 URL15 <http://www.promega.com>
 URL16 <http://www.caos.kun.nl>

Note: URLs can be unstable owing to the fluid nature of websites and corporate changes. Therefore the availability/accessibility of these references cannot be guaranteed.

© 1999 by the copyright holder of this paper. This paper may not be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or by any information storage and retrieval system, without permission in writing from the copyright holder.

XII

Abbreviations - Conventions
Definitions - Glossary - Terms

aa	amino acid
agenesis	absence of part of the body caused by an absent anlage
alae nasi	wings of the nose
aplasia	absence of part of the body caused by the failure of an anlage to develop
ARVCF	armadillo repeat gene deleted in velocardiofacial syndrome
association	a random occurrence in two or more individuals of multiple anomalies not known to be a polytopic field defect, sequence, or syndrome
ataxic	uncoordinated
atresia	congenital absence or closure of a normal body orifice or tubular organ
atresia of choanae	bony or membranous occlusion of the opening between nasal cavity and nasopharynx
atrophy	decrease in size of tissue or organ caused by decrease in cell size and/or number
auricles	external parts of external ears
ba1 [2, 3, ...], BA1 [2, 3, ...]	branchial arch 1 [2, 3, ...]
BEC1	see "CLDN5"
bp	basepair
CA	see "coarctation of the aorta"
CAFI	see "CHAFIB"
CAF-1	see "CHAFIB"
CAF1A	see "CHAFIB"
CAF1P60	see "CHAFIB"
canthus	angle of palpebral fissures
CDC45L	CDC45 (cell division cycle 45, S. cerevisiae, homolog)-like
CDC45L2	see "CDC45L"
CHAFIB	chromatin assembly factor 1, subunit B (p60)
chondrodysplasia punctata	stippled appearance of the epiphyses on X-ray images
clinodactyly	permanent lateral or medial deviation or deflection of one or more fingers
CLTCLI	clathrin, heavy polypeptide-like 1
CLDN5	claudin 5 (transmembrane protein deleted in velocardiofacial syndrome)
CPETRL1	see "CLDN5"
coarctation of the aorta	anomalous constriction in the aortic wall close to the ductus arteriosus
coloboma	partial closure of iris resulting in "keyhole" iris
cyctopia	one-eyedness
deformation	an abnormal form, shape or position of a part of the body caused by mechanical forces
dextrorotation of aorta	aorta connected to right ventricle
DGA	D:George Anomaly
DGCR	D:George Critical Region; D:George Chromosomal Region
DGS	D:George Syndrome
disease	condition of known cause in which there is progression and deterioration with time
disruption	a morphologic defect of an organ, part of an organ, or larger region of the body resulting from the extrinsic break-down of, or an interference with, an originally normal developmental process
DORV	see "double outlet right ventricle"
double inlet left ventricle	both left and right atrium connects to the left ventricle
double outlet right ventricle	both the pulmonary trunk and the aorta are connected to the right ventricle
dpc	days post coitum
dysphagia	difficulty swallowing
dysplasia	an abnormal organization of cells into tissue(s) and its morphological result(s); the process and consequence of dyshistogenesis
dystopia canthorum	fusion of inner eye-lids
ECM	extra cellular matrix
epicanthal folds	vertical folds left and right of the nose cover the fissures between eye-lids
epiphyses	ends of long bones
GP1BB	glycoprotein 1b (platelet), beta polypeptide
GSCL	gooseco'd-like
HS	see "PNUTLI"
HCDCREL-1	see "PNUTLI"
helices	posterior and superior free margin of the pinna of the ear
HIRA	HIR (histone cell cycle regulation defective, S. cerevisiae) homolog A
hyperplasia	overdevelopment of an organism, organ, or tissue resulting from a increased number of cells
hypertelorism	larger than normal distance between pupils
hypertrophy	increase in size of cells, tissue, or organ
hypoplasia	underdevelopment of an organism, organ, or tissue resulting from a decreased number of cells
hypospadias	male urethra opens ectopically between the glands and the scrotum in the ventral side of the penis due to ventral midline closure defect
hypotonia	reduced tone of skeletal muscles; reduced resistance to stretching
hypotrophy	decrease in size of cells, tissue, or organ
ichthyosis	scaling of skin

malformation	a morphologic defect of an organ, part of an organ, or larger region of the body resulting from an intrinsically abnormal developmental process
micrognathia	small chin
PA	see " pulmonary atresia "
palpebral	pertaining to eyelids
patent	open, unobstructed, not closed
patent ductus arteriosus	presence of the connection between left pulmonary artery and descending aorta
PDA	see " patent ductus arteriosus "
persistent truncus arteriosus	the two great arteries are connected to the ventricles as one vessel with one valve
philtrum	indented area between nose and upper lip
plnna	the projecting part of the ear lying outside of the head (also: auricles)
PNUTLI	Peanut (Drosophila)-like I
polydactyly	extra digits
polytopic field defect	a pattern of anomalies derived from the disturbance of a single developmental field
PS	see " pulmonary stenosis "
PTA	see " persistent truncus arteriosus "
pulmonary atresia	narrowing or obstruction of the opening between pulmonary artery and right ventricle
pulmonary stenosis	narrowing of opening between pulmonary and right ventricle
pulmonary valve stenosis	narrowing of opening between pulmonary and right ventricle at the level of the valves
PVS	see " pulmonary valve stenosis "
r1 [2, 3, ...]	rhombomere 1 [2, 3, ...]
RAA	see " right aortic arch "
RANBP1	RAN binding protein 1
retrognathia	withdrawn chin
right aortic arch	right aortic arch in stead of normal left arch
sequence	(clinical) a pattern of multiple anomalies derived from a single known or presumed prior anomaly or mechanical factor
SLC20A3	see " SLC25A1 "
SLC25A1	solute carrier family 25 (mitochondrial carrier, citrate transporter), member 1
STK22B	serine/threonine kinase 22B (spermiogenesis associated)
stridor	high pitched respiratory sound
syndactyly	fused digits
syndrome	a pattern of multiple anomalies thought to be pathogenetically related and not known to represent a single sequence or a polytopic field defect
syntropy	see " association "
TBX1	T-box 1
tetralogy of Fallot	combination of PS, VSD, dextroposition of the aorta and right ventricular hypertrophy, often combined with PA
TMVCF	see " CLDN5 "
TOF	see " tetralogy of Fallot "
transposition of great arteries	pulmonary connects to left ventricle and aorta to right ventricle
TSSK2	see " STK22B "
TUPLE1	see " HIRA "
UDFD1L	ubiquitin fusion degradation 1-like
VCFS	Ve/o Cardio Facial Syndrome
ventricular septum defect	persistent opening between left and right ventricle
VSD	see " ventricular septum defect "
WS1 [2,3,4]	Waardenburg Syndrome type 1 [2,3,4]

GENE	human gene or transcript
<i>Gene</i>	mouse or other non-human gene or transcript
PROTEIN	human protein
<i>Protein</i>	mouse or other non-human protein
P12345, Q67890	protein sequence database (SwissProt, TrEMBL) accession numbers
X12345, D67890, Z12345, AC678901	nucleotide sequence database (EMBL, GenBank) accession numbers
PIR-A12345	PIR protein database accession number
gdb 1234567	Genome DataBase identifier
E9.5	embryonic day 9.5 (9.5 dpc)

The following sources proved helpful in collating this data:

Spranger et al. 1982
 Dorland 1994
 URL3

XIII

Summary & Samenvatting

XIII 1 SUMMARY

In this thesis I describe our efforts to lift a tip of the veil covering the molecular genetics of the 22q11 chromosomal region associated with CATCH22. Syndromes that collectively form the CATCH22 or 22q11 Deletion Syndrome are characterized by their association with a deletion in chromosome 22q11 and a relation between phenotype and cranial neural crest. Survivability ranges from fatal (perinatal death) to eminently survivable or treatable. Accordingly, the clinical phenotypes are very variable: from mild craniofacial deformities to absence of thymus and severe heart conditions. Affected structures include, but are not limited to, thymus, thyroid, parathyroid, cardiac outflow tract, aortic arch derivatives, craniofacial structures, and inner and outer ears. The instance of all these structures being derived from, or dependent on, the embryonic neural crest, and being associated, in these syndromes, with deletion of part of chromosome 22q11 is very suggestive of one or more genes on or around 22q11 being involved in some aspect of neural crest development. Neural crest is a neuroectoderm derived population of migratory cells that contribute to many different types of tissues and organs, from enteric and parasympathetic ganglia to facial bones and cardiac septa. Though the genetics of neural crest development is being studied by many research groups employing techniques like transplantation, antisense treatment, knock-out transgenics and in vitro culture, much remains to be discovered. Many genes that do or could play a role in crest development have been identified and include genes as diverse as genes coding for homeobox, extracellular matrix, transmembrane receptor and extra-cellular signalling proteins. In order to find the 22q11 gene(s) contributing to cranial crest development we mapped genomic probes and markers to the deletion. Thus we could roughly determine the commonly deleted region and gain starting points for the construction of genomic contigs. Genomic sequences of these and additional contigs were analyzed for the presence of genes and features associated with genes. So far, the resulting map shows twentyseven genes (some putative) in the 1.2 Mb sequence analyzed. More than half of these

genes were unknown or not known to locate to chromosome 22 at the time. Gene and exon prediction programs were indispensable for supporting putative genes and were shown to be quite accurate in predicting and extending putative genes in those cases where the gene was identified and confirmed experimentally later. A large area in the sequence appears remarkably devoid of genes. One or more as yet unidentified (large) genes with low and/or localized expression may be located here. The nature of the gene products found in 22q11 is variable, without obvious kinship between the genes. I found an enzyme (MORF), putative nuclear regulators (RANBP1, HIRA), an ECM related factor (TOMCAT) and a homeobox protein (GSCL), amongst others. As one of the first and best candidate genes, *HIRA* was singled out for detailed analysis of expression and gene structure. It is expressed ubiquitously, but with enhanced expression in the branchial arches. Possible splice variants were identified in Northern blot essays, RNase protection assays and cDNA clone analysis. *HIRA* is a protein similar to yeast histone gene regulators Hir1p and 2p and was found to have homologs (orthologs?) in *C. elegans* and now also (by others) in *Drosophila*, *Fugu* and *Xenopus*. *HIRA* is part of the larger (and very diverse) family of WD40 repeat domain proteins. Recent experiments suggest that, as we indicated, the protein is involved in some form of chromatin interaction and regulation.

Recently one of the many 22q11 genes, *UFDIL*, has been put forward as a very strong candidate gene on account of its expression pattern and role in heart development. The variable nature of CATCH22 phenotypes and genotypes warrants further research into the role of other genes in 22q11 that could also be etiologic, either alone or in cooperation with *UFDIL* or other genes. *HIRA* and *GSCL*, amongst others, are strong contenders that are currently being studied through mutational analysis, knock-out and transgenesis. Future research should not only concentrate on studying 22q11 and its genes, but also explore the parallels between face and limb development and the currently largely neglected links between CATCH22 and chromosomes other than 22 (most notably 10).

In dit proefschrift beschrijf ik ons onderzoek dat ten doel heeft een tipje van de sluier op te lichten van de moleculaire genetica van het 22q11 chromosomale gebied dat met CATCH22 is geassocieerd. De syndromen die samen het CATCH22 of 22q11 Deletion Syndrome vormen worden gekarakteriseerd door de associatie met een deletie in chromosoom 22q11 en de relatie tussen fenotype en craniale neuuraallijst. Overleefbaarheid varieert van niet (overlijden rond geboorte) tot aan zeer (behandelbaar of geen noemenswaardige/levensbedreigende handicap). Het klinische fenotype is overeenkomstig variabel: van milde craniofaciale afwijkingen tot afwezigheid van thymus en ernstige hartafwijkingen. Aangedane structuren omvatten ondermeer thymus, schildklier, bijschildklier, uitstroomgebied van het hart, aortaboog afgeleiden, craniofaciale structuren, en binnenoer en oorschelp. Dat al deze structuren afstammen van de embryonale neuuraallijst, of er van afhankelijk zijn, en in deze syndromen met een deletie van een deel van chromosoom 22q11 zijn geassocieerd, suggereert dat een of meerdere genen in of in de buurt van 22q11 bij aspecten van neuuraallijst ontwikkeling betrokken zijn. Neuuraallijst is een van het neurectoderm afgeleide populatie van migrerende cellen welke bijdragen aan een reeks verschillende soorten weefsels en organen, van darm en parasympatische ganglia tot aan aangezichtsboten en hartmembranen. Hoewel de genetica van neuuraallijst ontwikkeling met behulp van technieken als transplantatie, anti-sense behandeling, knock-out transgenese en in vitro kweek door vele onderzoeksgroepen wordt onderzocht blijven er nog veel vragen open. Er zijn al talloze genen geïdentificeerd die zeker of mogelijk betrokken zijn bij neuuraallijst ontwikkeling, als divers als genen die coderen voor homeobox, extracellulaire matrix, transmembraan receptor eiwitten en extracellulaire signaal eiwitten. Teneinde de 22q11 genen te vinden die bijdragen aan neuuraallijst ontwikkeling, hebben wij genomische probes en markers in de deletie in kaart gebracht. Hierdoor konden wij het meest-gedeleteerde gebied van 22q11 bepalen en startpunten voor de constructie van genomische contigs verkrijgen. Genomische sequenties van deze en additionele contigs zijn geanalyseerd voor de aanwezigheid van genen en eigenschappen geassocieerd met genen. De kaart van het geanalyseerde gebied toont zevenentwintig genen (sommige daarvan potentiële genen) in 1.2 Mb sequentie. Meer dan de helft van deze genen waren onbekend of niet bekend dat ze op 22q11 lagen ten tijde van de analyse. Gen- en exonvoorspellingsprogramma's bleken onmisbaar voor het ondersteunen van potentiële genen en het blijkt dat deze programma's nauwkeurig potentiële genen voorspelden en uitbreidden waar deze genen later experimenteel werd geïdentificeerd en bevestigd. Een groot gebied in de geanalyseerde sequentie lijkt geen genen te bevatten. Een of meerdere nog te identificeren (grote) genen met lage en/of gelocaliseerde expressie kunnen hier aanwezig zijn. De aard van de genproducten die in 22q11 gevonden worden is zeer variabel, zonder enig duidelijk verband tussen de eiwitten. Onder andere identificeerde ik genen voor een

enzym (MORF), mogelijke nucleaire regulatoren (RANBP1, HIRA), een ECM geassocieerde factor (TOMCAT) en een homeobox protein (GSCL). Als een van de eerste en beste kandidaatgenen werd HIRA gekozen voor gedetailleerde expressie en genstructuur analyse. Het gen komt vrij algemeen tot expressie in het embryo, maar de kiewbogen tonen een hoger dan gemiddeld expressieniveau. Splice varianten van het mRNA werden geïdentificeerd met behulp van Northern assays, RNase protection assays en cDNA kloon analyse. HIRA is een eiwit dat lijkt op gist histon gen regulators Hir1p en Hir2p en een homoloog (ortholoog?) werd gevonden in de worm *C. elegans*. Recentelijk zijn door anderen ook orthologen gevonden in *Drosophila*, *Fugu* en *Xenopus*. HIRA maakt deel uit van een grotere (en heel heterogene) familie van WD40 repeat domein eiwitten. Recente experimenten suggereren dat, zoals wij al aangaven, het eiwit betrokken is bij chromatine interactie en regulatie.

Recentelijk is een van de vele 22q11 genen, *UFD1L*, naar voren gebracht als een sterk kandidaatgen vanwege het expressiepatroon en de rol in hartontwikkeling. Het feit dat CATCH22 fenotypen en genotypen zo variabel zijn maakt dat wij andere genen niet moeten uitsluiten van verder onderzoek. Andere genen, alleen of in combinatie met *UFD1L* of andere genen, kunnen ook bijdragen aan CATCH22 etiologie. *HIRA* en *GSCL*, maar ook andere genen, worden op dit moment onderworpen aan mutatie-analyse, knock-out experimenten en transgenese teneinde hun rol nader te kunnen bepalen. Toekomstig onderzoek zou zich niet alleen op 22q11 en de daarin aanwezige genen moeten concentreren, maar ook de parallellen tussen ledemaat- en aangezichtsontwikkeling en de op het moment ondergewaardeerde connecties tussen CATCH22 en andere chromosomen (met name chromosoom 10) moeten exploreren.

XIV

Curriculum Vitae

De schrijver van dit proefschrift werd op de eenentwintigste februari van het jaar negentienvijfenzestig geboren in de stad Utrecht. Na het behalen van het VWO diploma (keuzerichting exacte vakken) aan de Rijksscholengemeenschap F.A. Minkema te Woerden in 1983 begon hij aan de studie Biologie aan de Rijks Universiteit Utrecht. Gedurende de periode 1986-1987 werkte hij aan een project (rol van mRNA stabiliteit in de celcyclus regulatie van histon genen; verschil in mRNA interacties met het cytoskelet tussen membraangebonden en niet-membraangebonden mRNAs) aan the Department of Biochemistry and Molecular Biology, University of Florida College of Medicine, Gainesville, Florida, onder auspiciën van Prof. Dr G. Stein en Dr J. Stein in Gainesville en Prof. Dr H.O. Voorma in Utrecht. Gedurende de periode 1987-1988 werd de studie voor ruim een jaar onderbroken voor de vervulling van de militaire dienstplicht (hofmeester in NATO basis te Hohne, West-Duitsland). Na deze onderbreking werd voor de periode 1988-1989 een project bewerkt (*in situ* hybridizatie van *in vitro* gekweekte rat calvarium osteoblasten ten behoeve van onderzoek naar ontwikkelingsgeruleerde/-regulerende botspecifieke genen; electronenmicroscopisch onderzoek aan de door deze cellen afgescheiden collageen matrix) aan the Department of Cell Biology, University of Massachusetts Medical School, Worcester, Massachusetts, onder auspiciën van Prof. Dr G. Stein in Worcester en Prof. Dr J.A.M. v.d. Biggelaar in Utrecht. Vervolgens werd in Utrecht in Prof. Dr v.d. Biggelaar's Projectgroep Experimentele Embryologie, Vakgroep Experimentele Dierkunde, een project bewerkt (*in situ* hybridizatie van *Patella vulgata* ten behoeve van onderzoek naar ontwikkelingsregulatie van en door genen) waarna de studie in 1990 afgesloten werd met een doctoraal diploma Biologie met als specialisatie moleculaire celbiologie, ontwikkelingsbiologie en genetica. Van 1991 tot en met 1997 bewerkte hij een promotieproject onder begeleiding van Dr J.H.C. Meijers, Prof. Dr F. Grosveld en Prof. Dr J. Molenaar aan de vakgroep Celbiologie en Genetica van de faculteit Geneeskunde en Gezondheidswetenschappen van de Erasmus Universiteit Rotterdam in samenwerking met de afdeling Kinderheelkunde van het Sophia Kinderziekenhuis. Dit promotieproject, over HIRA, Chromosoom 22q11 en CATCH22, culmineerde in het proefschrift dat nu voor u ligt.

Sinds 1997 is de schrijver werkzaam als Senior Computer Biologist aan the Informatics Department, Sanger Centre, Hinxton, United Kingdom, als deel van het Human Genome Project en Mouse Genome Project.



Publications

- Dunham I, Hunt AR, Collins JE, Bruskewich R, Beare DM, Clamp M, Smink LJ, Ainscough R, Almeida JP, Babbage A, Bagguley C, Baily J, Barlow K, Bates KN, Beasley O, Bird CP, Blakey S, Bridgeman AM, Buck D, Burgess J, Burrill WD, Burton J, Carder C, Carter NP, Chen Y, Clark G, Clegg SM, Cobley V, Cole CG, Collier RE, Connor RE, Conroy D, Corby N, Coville GJ, Cox AV, Davis J, Dawson E, Dharni PD, Dockree C, Dodsworth SJ, Durbin RM, Ellington A, Evans KL, Fey JM, Fleming K, French L, Garner AA, Gilbert JGR, Goward ME, Grafham D, Griffiths MN, Hall C, Hall R, Hall-Tamlyn G, Heathcote RW, Ho S, Holmes S, Hunt SE, Jones MC, Kimberley A, King A, Laird GK, Langford CF, Leversha MA, Lloyd C, Lloyd DM, Martyn ID, Mashregi-Mohammadi M, Matthews L, McCann OT, McClay J, McLaren S, McMurray AA, Milne SA, Mortimore BJ, Odell CN, Pavitt R, Pearce AV, Pearson D, Phillimore BJ, Phillips SH, Plumb RW, Ramsay H, Ramsey Y, Rogers L, Ross MT, Scott CE, Sehra HK, Skuce CD, Smalley S, Smith ML, Soderlund C, Spragon L, Steward CA, Sulston JE, Swann RM, Vaudin M, Wall M, Wallis JM, Whiteley MN, Willey D, Williams L, Williams S, Williamson H, Wilmer TE, **Wilming L**, Wright CL, Hubbard T, Bentley DR, Beck S, Rogers J, Shimizu N, Minoshima S, Kawasaki K, Sasaki T, Asakawa S, Kudoh J, Shintani A, Shibuya K, Yoshizaki Y, Aoki N, Mitsuyama S, Roe BA, Chen F, Chu L, Crabtree J, Deschamps S, Do A, Do T, Dorman A, Fang F, Fu Y, Hu P, Hua A, Kenton S, Lai H, Lao H, Lewis J, Lewis S, Lin S-P, Loh P, Malaj E, Nguyen T, Pan H, Phan S, Qi S, Qian Y, Ray L, Ren Q, Shaull S, Sloan D, Song L, Wang Q, Wang Y, Wang Z, White J, Willingham D, Wu H, Yao Z, Zhan M, Zhang G, Chissoe S, Murray J, Miller N, Minx P, Fulton R, Johnson D, Bemis G, Bentley D, Bradshaw H, Bourne S, Cordes M, Du Z, Fulton L, Goela D, Graves T, Hawkins J, Hinds K, Kemp K, Latreille P, Layman D, Ozersky P, Rohlfing T, Scheet P, Walker C, Wamsley P, Wohldmann P, Pepin K, Nelson J, Korl I, Bedell J, Hillier L, Mardis E, Waterston R, Wilson R, Emanuel BS, Shaikh T, Kurahashi H, Saitta S, Budarf ML, McDermid HE, Johnston A, Wong ACC, Morrow BE, Edelmann L, Kim U-J, Shizuya H, Simon MI, Dumanski JP, Peyrard JP, Kedra D, Seroussi E, Fransson I, Tapia I, Bruder CEG, O'Brien KP

The DNA sequence of human chromosome 22.

Nature (1999) **402**: 489-495

- **Wilming LG**, Snoeren CA, van Rijswijk A, Grosveld F, Meijers C
The murine homologue of HIRA, a DiGeorge syndrome candidate gene, is expressed in embryonic structures affected in human CATCH22 patients.
Hum Mol Genet (1997) **6**: 247-258
- Mulder MP, Wilke M, Langeveld A, **Wilming LG**, Hagemeijer A, van Drunen E, Zwarthoff EC, Riegman PH, Deelen WH, van den Ouweland AMW, Halley DJJ, Meijers C
Positional mapping of loci in the DiGeorge critical region at chromosome 22q11 using a new marker (D22S183).
Hum Genet (1995) **96**: 133-141
- Baarends WM, van Helmond MJL, Post M, van der Schoot PJCM, Hoogerbrugge JW, de Winter JP, Uilenbroek JTJ, Karels B, **Wilming LG**, Meijers JHC, Themmen APN, Grootegoed JA
A novel member of the transmembrane serine/threonine kinase receptor family is specifically expressed in the gonads and in mesenchymal cells adjacent to the müllerian duct.
Development (1994) **120**: 189-197
- Pockwinse SM, **Wilming LG**, Conlon DM, Stein GS, Lian JB
Expression of cell growth and bone specific genes at single cell resolution during development of bone tissue-like organization in primary osteoblast cultures.
J Cell Biochem (1992) **49**: 310-323
- Zambetti G, **Wilming L**, Fey EG, Penman S, Stein J, Stein G
Differential association of membrane-bound and non-membrane-bound polysomes with the cytoskeleton.
Exp Cell Res (1990) **191**: 246-255
- Owen TA, Aronow M, Shalhoub V, Barone LM, **Wilming L**, Tassinari MS, Kennedy MB, Pockwinse S, Lian JB, Stein GS
Progressive development of the rat osteoblast phenotype in vitro: reciprocal relationships in expression of genes associated with osteoblast proliferation and differentiation during formation of the bone extracellular matrix.
J Cell Physiol (1990) **143**: 420-430

XVI

Dankwoord

In tegenstelling tot wat het geval was vele eeuwen geleden gedurende het tijdperk van de Uomo Universalis, toen individuen alle velden van wetenschap konden bestuderen en monografiën erover schreven, is dit proefschrift natuurlijk niet het resultaat van een persoon alleen. Het is tot stand gekomen dankzij het werk van velen. Ik noem met name Dr Carel Meijers en Prof. Dr Jan Molenaar voor het in mij gestelde vertrouwen, zonder welke ik het project nooit had kunnen beginnen en afronden. Carel, bedankt voor je geduld toen proefschrift-deadline na deadline verstreek en voor de richting die je gegeven hebt tijdens de fasen in het onderzoek dat ik die verloren had. Ik waardeer de vele discussies die we gehad hebben en benijd je schijnbaar bodemloze geheugen waar citaten, feiten en referenties uit komen stromen als water uit een lek vat. Ik ben natuurlijk veel dank verschuldigd aan Sylvia Snoeren-Soffers, die mij bijstond bij vele experimenten en wiens transformatie van Soffers naar Snoeren mij een slapeloze nacht bezorgde (Sylvia, je huwelijkscadeautje was pas luttele uren voor overhandiging uit de oven gekomen voor het bakken van de verf). Ik kijk met plezier terug naar de goede tijd die ik heb gehad op het lab. Hiervoor dank ik ook Marjo Peters, die mij geleerd heeft op kippen embryo's te opereren. Zolang ik maar niet haar instrumenten gebruikte, want ik maakte die in no-time bot. Bedankt ook voor de vele barbecues Marjo en Wil, en dat er nog veel mogen volgen! Martina, ich danke auch dich, jij bent met de cosmiden begonnen voor ik het overnam. Ik denk graag terug aan de heerlijke diners die ik gehad heb met jou en Bob. Jullie hebben mij ingewijd in de Japanse keuken, dat zal ik nooit vergeten. Marie-Josée, ook deel van de langstlevende labpopulatie, maar wel van "de andere kant van de labtafel": je hebt geprobeerd mij enigzins in bedwang te houden als de labconversaties paden betraden die niet voor drukwerk geschikt zijn. Sorry dat dat meestal niet lukte. Aan jouw kant waren ook Sylvia D., An en Dirk-Jan. Ook jullie bedankt voor de vele jaren discussies en conversaties over de meest rare onderwerpen. An, bedankt voor het FISHen en je hulp bij de receptie.

Mensen kwamen en gingen in het lab dus er zijn nog vele die ik wil bedanken voor hun aanwezigheid. Guy, Vincent, thanks for the great fun; Anja, ik mis iemand om te pesten sinds ik jou niet meer op de stoel naast me heb; Sandra, il tuo temperamento infuocato colora la vita; Ragonda, bedankt voor je experimentele ondersteuning en de tijd die we samen hadden; Nynke, met jou als overbuurvrouw had ik altijd een mooi uitzicht; Frans: praat je nog steeds zoveel? Iris, Theo, Gerard, Robert, Niels, Arnout, Clara, Erik, Sjaak, Maarten M. en al die ik vergeten ben: jullie hebben voor kortere of langere tijd het lab kleur gegeven. Ik mis trouwens al die taarten en gebakjes die talloze verjaardagen of andere vieringen vergezelden.

Verder bedank ik de fotografen Tom, Mirko en Ruud voor hun hulp bij publicaties, Marike voor haar secretariële ondersteuning, Rein, Mieke en Melle voor hun assistentie bij mijn taak als "consumables bestelmeester". Dank aan Dr Marja Wessels het ter beschikking stellen van de patiëntenfoto's en de daarop afgebeelde mevrouw voor de toestemming voor publicatie ervan. Ton Verkerk, je bent onmisbaar geweest voor de sequentie analyse en hebt mij geïntroduceerd in de arcane wereld van Unix. In niet geringe mate dankzij jou ben ik beland waar ik nu ben. Veel dank.

Iwould like to thank Dr Peter Scambler for our collaboration on CATCH22. I also thank Prof. Dr Rolf Zeller and Prof. Dr Elaine Dzierzak for reviewing the manuscript. Ik ben dank verschuldigd aan commissielid Prof. Dr Ben Oostra voor de tijd die hij genomen heeft om mijn manuscript te lezen. Dank vooral ook aan mijn twee promotoren Prof. Dr Frank Grosveld en Prof. Dr Dick Tibboel die de taak op zich hebben genomen de procedures in goede banen te leiden en het manuscript meerdere malen te lezen. Frank, bedankt voor het introduceren van de Engelse traditie van Christmas pantomimes. Ik heb met erg veel plezier bijgedragen aan een paar ervan en het doet mij deugd dat de props iedereen nog steeds elke morgen begroeten als ze uit de lift stappen. Now we're on the subject: Dubi, thanks for making the statue that was given to me as the first price in the first Christmas pantomime video competition, and for the various diners in exchange for my technical assistance.

Zeer veel dank aan paranymphen Maarten, Hetty en (weer) Sylvia. Zonder jullie waren het proefschrift en de bijbehorende festiviteiten er nooit geweest. Een lange-afstands promovendus heeft een afstandsbediening als jullie hard nodig. Hetty, je hebt je kwaliteiten als organizer kunnen botvieren in het organiseren van de festiviteiten; Maarten, je kennis op het gebied van graphica en elektronisch publiceren waren van onschatbare waarde; Sylvia, je diende als mijn geheugen waar het mijne faalde.

Als laatste bedank ik mijn ouders die mij altijd hebben bijgestaan in de keuzes die ik maakte en mij hebben aangespoord het beste eruit te halen.



Colophon

Graphics:

Adobe Illustrator, Adobe Photoshop, Corel Draw, Deneba Canvas, Macromedia Freehand, Metacreations Raydream 3D.

Text:

Nisus Writer.

Layout:

Quark XPress.

Hardware:

Apple Macintosh IIci/Sonnet 040/80, Quadra 700/PPC601/50, Quadra 950/Sonnet PPC601/100, PM9600 200MHz, PB1400c/NewerTech PPC750/250.