Department of Molecular Medicine, Karolinska Institutet,
Stockholm, Sweden

CHARACTERIZATION OF HUMAN CHROMOSOME 22: CLONING OF BREAKPOINTS OF THE CONSTITUTIONAL TRANSLOCATION t(11;22)(q23;q11) AND DETECTION OF SMALL CONSTITUTIONAL DELETIONS BY MICROARRAY CGH.

Isabel Tapia Páez

Stockholm 2003

Published and printed by Karolinska University Press
Box 200, SE-171 77 Stockholm, Sweden

ISBN 91-7349-505-0

I find a great part of the information I have was acquired by looking up something and finding something else on the way.

Franklin P. Adams

To Micaela

ABSTRACT

Chromosome 22 is the second smallest human chromosome, composing approximately 1.5% of the genome. The short arm of this acrocentric chromosome harbors ribosomal genes and the long arm contains protein coding genes. This chromosome is gene-rich in comparison to the majority of other chromosomes, containing approximately 600 characterized genes to date. Many of these are involved in the etiology of a wide spectrum of diseases such as congenital and psychiatric disorders as well as cancers.

The constitutional translocation t(11;22) is the most common reciprocal translocation in humans. This translocation is often found in families but can also occur *de novo*. Translocation carriers are normal and usually become diagnosed in connection with infertility problems or a birth of a genetically unbalanced child. In addition, an increased risk to breast cancer has been reported in some carriers, which suggests that the translocation might have an effect on a gene(s) involved in the etiology of breast cancer. We characterized the breakpoints of this translocation and found that the breakpoint region on chromosome 22 lies within an unclonable gap. The breakpoint on chromosome 11 is also located within an unstable region, as all BACs containing this segment are rearranged. We identified one BAC from chromosome 11 spanning the translocation breakpoint and two BAC clones from chromosome 22, which contain sequences similar/identical to the sequences mediating the translocations breakpoints on chromosome 22. A cosmid library from one translocation carrier was also constructed and chimeric cosmids from both derivative chromosomes were isolated. Their analysis revealed that no gene(s) seems to be disrupted by the translocation breakpoints. We also show that the breakpoints on both chromosomes occur at the tip of hairpins, which are formed due to the presence of long inverted repeats/palindromes. The formation of these structures is the likely reason behind "unclonability" of this region on chromosome 22 and the instability of BACs derived from chromosome 11. Furthermore, based on fiber-FISH experiments we conclude that the breakpoints of the translocations are highly conserved among carriers.

The second aspect of the thesis is related to detection of micro-deletions and micro-gains, which cause a large number of genetic disorders. In order to improve the detection of such rearrangements, we applied and further developed the microarray-CGH methodology. We constructed three microarrays: one covering 7 Mb region in the vicinity of the *NF2* gene in 22q12; the second is a full coverage chromosome 22 array; and the third is an array covering 6 Mb from the 22q11 region, including the typically deleted region in DiGeorge/Velo-Cardio-Facial syndrome. The latter region is particularly challenging, due to the presence of low copy repeats, high content of common repeats and unclonable sequences. Three types of targets were used in the arrays: ***i)*** genomic clones; ***ii)*** non-redundant, repeat-free pools of genomic DNA amplified by PCR; and ***iii)*** cDNA-based targets, single as well as in pools. We used the arrays to study neurofibromatosis type 2, acral melanoma, dermatofibrosarcoma, and DiGeorge/Velo-Cardio-Facial syndrome. We were able to detect homozygous/heterozygous deletions, amplifications, *IGLV/IGLC* locus instability and the breakpoints of an imbalanced translocation. Using the novel approach with repeat-free, PCR-generated sequences, we detected heterozygous deletions using as little as 11.5 kb of genomic target sequence. We conclude that the array-CGH is a powerful method for the detection of gene-dosage imbalances. Our results also suggest that most, if not all, medically important segments of our genome will be accessible for analysis using high-resolution microarray-based CGH.

LIST OF PUBLICATIONS

I. Fine mapping of the constitutional translocation t(11;22)(q23;q11).
Tapia-Páez I, O'Brien KP, Kost-Alimova M, Sahlen S, Kedra D, Bruder CE, Andersson B, Roe BA, Hu P, Imreh S, Blennow E, Dumanski JP (2000) Human Genetics 106: 506-516.

II. The DNA sequence of human chromosome 22.
Dunham I, Shimizu N, Roe BA, Chissoe S, Hunt AR, Collins JE, Bruskiewich R, Beare DM, Clamp M, Smink LJ, Ainscough R, Almeida JP, Babbage A, Bagguley C, Bailey J, Barlow K, Bates KN, Beasley O, Bird CP, Blakey S, Bridgeman AM, Buck D, Burgess J, Burrill WD,et al, **Tapia I,** et al. (1999). Nature 402: 489-495

III. The position of t(11;22)(q23;q11) constitutional translocation breakpoints are conserved among its carriers.
Tapia-Páez I*, Kost-Alimova M*, Hu P, Limei Yang, Roe BA, Blennow E, Ludmila Fedorova, Imreh S, Dumanski JP (2001) Human Genetics 109: 167-177.

IV. High resolution deletion analysis of constitutional DNA from neurofibromatosis type 2 (NF2) patients using microarray-CGH.
Bruder CE, Hirvela C, **Tapia-Páez I**, Fransson I, Segraves R, Hamilton G, Zhang XX, Evans DG, Wallace AJ, Baser ME, Zucman-Rossi J, Hergersberg M, Boltshauser E, Papi L, Rouleau GA, Poptodorov G, Jordanova A, Rask-Andersen H, Kluwe L, Mautner V, Sainio M, Hung G, Mathiesen T, Moller C, Pulst SM, Harder H, Heiberg A, Honda M, Niimura M, Sahlen S, Blennow E, Albertson DG, Pinkel D, Dumanski JP (2001) Human Molecular Genetics 10: 271-282.

V. A full-coverage and high-resolution human chromosome 22 microarray; clinical and research applications.
Buckley P. G., Mantripragada K. K., Benetkiewicz M., **Tapia-Páez I**., Diaz De Stahl T., Rosenquist M., Ali H., Jarbo C., De Bustos C., Hirvela C., Sinder Wilen B., Fransson I., Thyr C., Johnsson B. I., Bruder C. E., Menzel U., Hergersberg M., Mandahl N., Blennow E., Wedell A., Beare D. M., Collins J. E., Dunham I., Albertson D., Pinkel D., Bastian B. C., Faruqi A. F., Lasken R. S., Ichimura K., Collins V. P., and Dumanski J. P. Human Molecular Genetics 11:3221-3229

VI. DNA copy number analysis of the typically deleted region in DiGeorge syndrome using array-CGH with genomic and PCR-based targets.
Tapia-Páez I*, Mantripragada KK*, Blennow E, Wedell A, Dumanski JP. Manuscript

*Authors contributed equally to this work.

1 CONTENTS

LIST OF ABBREVIATIONS / GLOSSARY

ABL	Abelson murine leukemia virus oncogene
ANILFR	Average Normalized Inter-Locus Fluorescence Ratio
BAC	Bacterial Artificial Chromosome
BCR	Breakpoint Cluster Region
Bp	Base pair(s)
CES	Cat eye syndrome
cDNA	Complementary DNA
CGH	Comparative genomic hybridization
Contig	Group of overlapping sequences. Also, a group of overlapping genomic clones can form a contig.
Cosmid	Artificial constructed vector, containing cos sites
ddNTP	Dideoxy nucleotide triphosphate
DFSP	Dermatofibrosarcoma protuberans
DGS/VCFS	DiGeorge/velocardiofacial syndrome
DNA	Deoxiribonucleic Acid
EST	Expressed Sequence Tag
FISH	Fluorescence *in-situ* Hybridisation
IGLC	Immunoglobulin lambda constant region
IGLV	Immunoglobulin lambda light chain, variable region genes
Kb	Kilobases
LCR	Low Copy Repeats
LOH	Loss of Heterozygosity
Mb	Megabases
NF1	Neurofibromatosis type 1
NF2	Neurofibromatosis type 2
PAC	P1 Artificial Chromosome
PCR	Polymerase Chain Reaction
STS	Sequence tagged site
TDR	Typically Deleted Region (referred to the common deleted region in DGS/VCFS)

2 INTRODUCTION

2.1 CHROMOSOME 22

Chromosome 22 is the second smallest human chromosome, consisting of 1.6-1.8% of the human genome. The short arm (22p) of this acrocentric chromosome harbors ribosomal genes, while the long arm (22q) contains the protein coding genes, and it is this region which has been sequenced. This chromosome is gene-rich in comparison to other chromosomes, with 16.3 genes per Mb. For comparison the least gene-rich chromosome 13 has five genes per Mb and the most gene-rich chromosome 19 has 23 genes per Mb.

Today, approximately 600 genes have been described mapping to this chromosome. Many of these are involved in the etiology of a wide spectrum of diseases. Examples include congenital developmental disorders as DiGeorge/velo-cardio-facial (DGS/VCFS) and cat-eye syndromes (CES), cancers in Ewing sarcoma, chronic myeloid leukemia, neurofibromatosis type 2, dermatofibrosarcoma protuberans, breast cancer, psychiatric disorders as schizophrenia and mental retardation (Figure 1). Therefore, chromosome 22 has caught the attention of many research groups.

Due to its medical importance, this chromosome has been in the forefront of research. First, a comprehensive linkage map over chromosome 22 was produced (Dumanski *et al.*, 1991), and later a YAC-based physical map was constructed (Collins *et al.*, 1995). It was the first chromosome sequenced (Dunham *et al.*, 1999) and recently we presented a full coverage chromosome 22 microarray for detection of gene dosage alterations (Buckley *et al.*, 2002).

Since chromosome 22 is one of the smallest chromosomes with a high medical importance it was natural that it became the first chapter released of the “book of life” (the sequence of the human genome). The sequencing of chromosome 22 was performed as part of an international effort in sequencing the entire human genome and it is a milestone in the understanding of our human biology. The sequencing report in 1999 provided the first view of the complex chromosomal landscapes found in the human genome. This sequence was estimated to cover 97% of 22q consisting of 12 contiguous segments covering 33.4 million bp separated by 11 gaps of known size. Despite all the efforts in closing the gaps, to date only one of these gaps in 22q11 has been closed by the Advanced Center for Genome Technology University of Oklahoma. One

interesting feature of chromosome 22 is the presence of Low Copy Repeats specific for this chromosome (LCR22) (Dunham *et al.*, 1999; Edelmann *et al.*, 1999a; Edelmann *et al.*, 1999b). These LCRs consist of blocks of tandem repeats containing copies of the γ-glutamyl transferase genes and the BCR-like genes. Details about the LCR22 are given in a separate chapter.

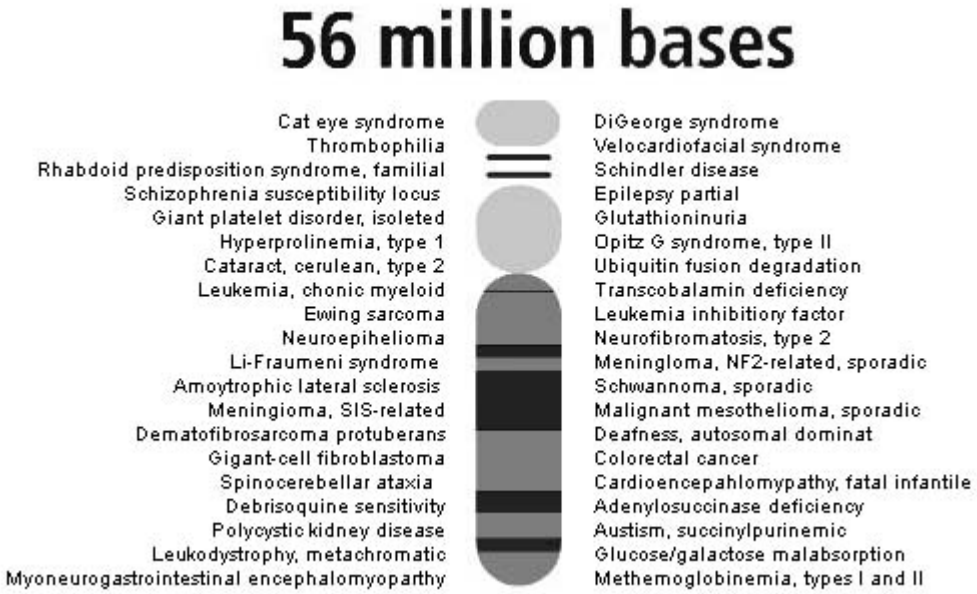


Figure 1. Chromosome 22 and list of diseases mapped to this chromosome.

2.2 CHROMOSOMAL REARRANGEMENTS

Chromosomal aberrations can be divided into two groups: numerical and structural. Numerical are those where the number of chromosomes is altered. If an entire set of chromosomes is involved it causes *polyploidy*, for example when two sperms fertilize a single egg, a triploid is formed. Polyploidy is not compatible with life in humans or animals, but is seen in plants. *Aneuploidy* is when single chromosomes are gained or lost. The only non-lethal monosomy in humans is monosomy X in Turner syndrome. Autosomal monosomies often lead to spontaneous abort early in pregnancy. Extra chromosomes are more common, trisomy 21 (Down syndrome) is the most common chromosomal abnormality seen in humans with an incidence of 1-2 in 1,000 live births. Other human trisomies are trisomy 13 (Patau syndrome) and trisomy 18 (Edwards syndrome).

The second group of chromosomal aberrations consist of changes in chromosomal structure. Structural abnormalities are common events and often lead to gains or losses of chromosomal segments causing increased or decreased expression of genes. The structural abnormalities are balanced or unbalanced depending on whether there is a loss or gain of genetic material or not.

In general, balanced structural abnormalities do not affect the phenotype but there are some exceptions (see translocations). A correct gene dosage is essential for normal development, and unbalanced chromosomal aberrations cause a wide range of developmental disorders, characterized by malformations, growth retardation and mental retardation. Another group of genes sensitive to gene dosage alterations are the genes related to tumorigenesis. Mutations causing loss of function may lead to the inactivation of tumor suppressor genes. The normal function of these is to inhibit the cell from excessive proliferation, thus, loss of function mutations result in the development of cancer. Mutations causing "gain of function" in oncogenes also result in development of tumors. The normal function of oncogenes is to promote cell proliferation. An increased dosage of oncogenes may cause excessive or inappropriate activation of cell proliferation.

The first event in the formation of structural chromosomal rearrangements is the breakage of DNA. Among the known causes of DNA breakage are ionizing radiation and chemical mutagens, followed by a deficiency in the mechanism of recombination between nonhomologous chromosomes or misrepair of the chromosome breaks.

A graphical view of structural chromosomal rearrangements is presented in Figure 2. Structural chromosomal abnormalities can be divided in:

(a) Deletions. A deletion is a loss of a chromosomal segment. Deletions caused by two breaks in the chromosome yield an interstitial deletion. If only one break occurs at the end of the chromosome it is a terminal deletion. Sometimes breaks occur at both ends of a chromosome, if these ends are joined they form a "ring chromosome", if the centromere is present it may pass through cell division. Deletions are always unbalanced as there is loss of genetic material. Sometimes these losses are referred to as "partial monosomy" for a specific segment.

(b) Duplications. This type of aberration also belongs to the category of unbalanced rearrangements as a segment of a chromosome is repeated. The segment can be repeated one or several times, the orientation of the repeat may be the same as the original segment (tandem repeat) or it may be in an opposite orientation (inverted duplication).

(c) Inversions. These are often balanced rearrangements as there is no loss of genetic material. Inversions are caused by two breaks in DNA, the chromosome pieces are repaired and joined in an inverted orientation. Paracentric inversions occur if the breaks are in the same arm, if the breaks are on either side of the centromere the inversion is pericentric.

(d) Translocations are chromosomal abnormalities which occur when chromosomes break and the fragments rejoin to different chromosomes than they originate from. Thus, there is an exchange of genetic material between chromosomes.

As the translocation t(11;22)(q23;q11) is one of the main topics of this thesis, more details about translocations will be discussed.

Chromosomal translocations may be balanced or unbalanced depending on whether there is loss or gain of material or not. Translocations are also divided in *reciprocal translocations* when there is an exchange of material between terminal segments from two chromosomes and *Robertsonian translocations*, when the breaks occur in the short arms of two acrocentric chromosomes, the centromeres from the two chromosomes are fused and a "new" large chromosome is formed. In Robertsonian translocations there is loss of material from the short arms but these translocations are considered balanced since the short arms of acrocentric chromosomes contain only ribosomal genes. Ribosomal genes are present in many copies in the short arms of the other acrocentric chromosomes and no phenotypic effects are observed from this loss.

Most often a balanced chromosomal translocation does not affect the phenotype as it may occur in regions were there are no genes. Sometimes, however, chromosomal translocations lead to the disruption of a particular gene or to the juxtaposition of two independent genes inducing inappropriate expression of the involved genes or the synthesis of a new "fusion protein". This phenomenon is seen in cancers when the translocation breakpoints affect oncogenes or tumor suppressor genes.

Some examples of translocations involved in cancers are:

- The t(9;22) involving the *ABL* oncogene from chromosome 9 and the *BCR* gene from chromosome 22. This translocation leads to the formation of the philadelphia chromosome (Ph[1]) containing a chimeric fusion protein *ABL/BCR*. This phenomenon is seen cytogenetically in over 90% of cases of chronic myeloid leukemia (CML).

- The t(8;14) juxtaposes the *MYC* oncogene from chromosome 8 with the immunoglobulin heavy chain (*IGH*) from chromosome 14, leading to over-expression of *MYC*. This translocation is seen in Burkitt´s lymphoma.

- In Dematofibrosarcoma protuberans (DFSP) the t(17;22)(q22;q13) fuses the plateled-derived growth factor B-chain gene (*PDGFB* proto-oncogene) from chromosome 22 with the collagen-type I alpha (*COL1A1*) gene from chromosome 17, resulting in a chimeric fusion protein

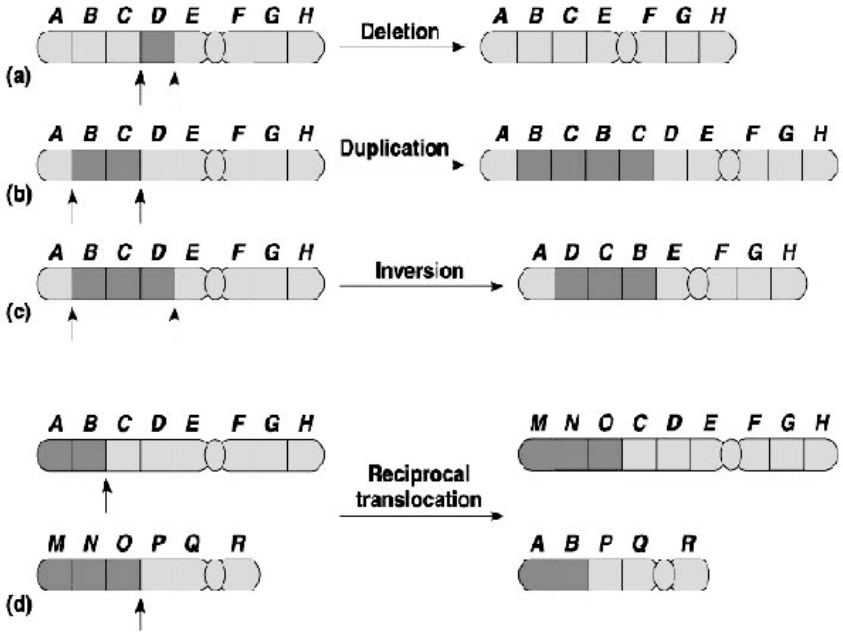


Figure 2. Examples of structural chromosomal abnormalities

2.3 METHODS FOR DETECTION OF CHROMOSOMAL ABERRATIONS

Genomic disorders represent an important entity among human genetic diseases. The detection of these has improved dramatically with the advances in molecular genetics. In 1960, chromosomes could be identified by their size and the position of the centromeres. In 1968 Caspersson (Caspersson *et al.*, 1968; Caspersson *et al.*, 1970) developed the Q-banding technique and in the 1970s other banding techniques such as C-, G-, R- and T-banding methods were developed. Banding techniques allow the visualization of differentially stained regions of a chromosome as a continuous series of light and dark bands that are, specific for each chromosome. These techniques allow a more accurate identification of each chromosome. However, the resolution of chromosome analysis is still low (approximately 1-3 Mb), and only large rearrangements can be detected. In the 1990s the FISH method became available, dramatically increasing the level of resolution for detection of chromosomal abnormalities (Trask, 1991; van Ommen *et al.*, 1995). The position of the rearrangement has to be known and a specific probe has to be available for this method. Today, the array-CGH method is emerging as a powerful tool in detecting micro-deletions and micro-duplications with high resolution on a large scale.

2.4 THE 22q11 REGION AND THE PRESENCE OF LOW COPY REPEATS SPECIFIC FOR CHROMOSOME 22 (LCR22)

The 22q11 region is involved in several chromosomal rearrangements that are associated with congenital and neoplastic diseases such as translocations, deletions and duplications.

Different recurrent translocations have breakpoints in 22q11, such as the t(9;22) associated with acute lymphocytic leukemia (ALL) and chronic myeloid leukemia (CML), the translocation t(8;22) associated with Burkitt´s lymphoma and the constitutional translocation t(11;22)(q23;q11). Constitutional micro-deletions and micro-duplications in this region cause DiGeorge (DGS)/velo-cardio-facial (VCFS) and cat eye (CES) syndromes, respectively. The accumulation of all these rearrangements in this particular region suggests, that this is a hotspot for chromosomal rearrangements in our genome.

Sequence analysis of the human genome has lead to the identification of segmental duplications, these can be either inter- or intrachromosomal. Since one of the topics of this thesis is specific intrachromosomal duplications, I will focus on these.

Today we know some chromosomal rearrangements associated with Low Copy Repeats (LCRs). For example, the repeats in chromosome 17p12 mediate interstitial duplications or deletions associated with the Charcot-Marie-Tooth disease type1A (CMT1A) and hereditary neuropathy with pressure palsies (HNPP) respectively (Chance *et al.*, 1994). Segmental duplications in 17p11.2 mediate deletions associated with Smith-Magenis syndrome (SMS) and its reciprocal interstitial duplication product (Chen *et al.*, 1997). LCRs also mediate deletions in chromosome 15q11 underlying Prader-Willy and Angelman syndromes (Christian *et al.*, 1999).

In chromosome 22, eight LCRs have been identified which are specific for this chromosome (Dunham *et al.*, 1999; Edelmann *et al.*, 1999a; Shaikh *et al.*, 2000). The structure and organization of these LCRs are very complex, comprising blocks of tandem repeats with the presence of truncated genes and pseudogenes like GGTL, BCRL, V7rel, KIAA0649L and others. Four of these LCRs (in this thesis LCR-A through LCR-D) are located within the 3Mb typically deleted in DGS/VCFS and the majority of the DGS/VCFS patients have proximal and distal endpoints of the deletion in these LCR. All LCRs have been fully sequenced except for LCR-B which is the LCR containing the breakpoints in chromosome 22 of the constitutional translocation t(11;22)(q23;q11).

FISH studies have shown that these LCR22 have arisen recently in evolution; they are not present in the mouse but they are present in primates suggesting an origin at least 40 million years ago (Shaikh *et al.*, 2000).

2.5 THE CONSTITUTIONAL TRANSLOCATION t(11;22)(q23;q11)

The constitutional translocation t(11;22)(q23;q11) is a recurrent reciprocal, balanced translocation, it is the most common translocation found in humans, being reported in more than 100 unrelated families (Fraccaro *et al.*, 1980; Iselius *et al.*, 1983; Zackai & Emanuel, 1980). This translocation is often found in families, but also occurs *de novo*. Translocation carriers have no clinical symptoms and are often detected when infertility problems and/or miscarriages arise. They also have an increased risk of having children with an unbalanced translocation, leading to the der(22) syndrome (Beedgen *et al.*, 1986). The main characteristics of children with der(22) syndrome are craniofacial abnormalities, heart defects and mental retardation resembling the DGS/VCFS syndrome spectrum (Funke *et al.*, 1999).

It has been reported that in a group of eight families carrying the constitutional translocation t(11;22)(q23;q11), there was an increased risk of developing breast cancer (Lindblom *et al.*, 1994). This suggested that there could be effects on a gene(s) responsible for the aetiology of breast cancer in this group of patients mediated by the translocation.

For this reason, substantial efforts were made from several groups in the attempt of cloning the t(11;22)(q23;q11) breakpoints (Edelmann *et al.*, 2000; Hill *et al.*, 2000; Shaikh *et al.*, 1999; Tapia-Paez *et al.*, 2001). The cloning of the translocation breakpoints has not been an easy task, mainly due to the presence of Low Copy Repeat sequences surrounding the translocation breakpoint in chromosome 22. There is also an unclonable gap in this region. One group suggested that the translocation was mediated by ALU repeats (Hill *et al.*, 2000), but this result was later disclaimed and proved to be a PCR artifact (Kurahashi *et al.*, 2000a).

Recently, the breakpoints of another constitutional translocation, t(17;22) have been cloned (Kurahashi *et al.*, 2003). The carrier of this translocation is a patient with NF1, the translocation disrupted the NF1 gene in 17q11 and the breakpoint in chromosome 22 is at the same position as in the t(11;22), suggesting that this region is a hot spot for other reciprocal translocations. The latter statement is also suggested by the fact that a high rate of *de novo* translocation t(11;22) was observed in sperm from normal males (Kurahashi & Emanuel, 2001).

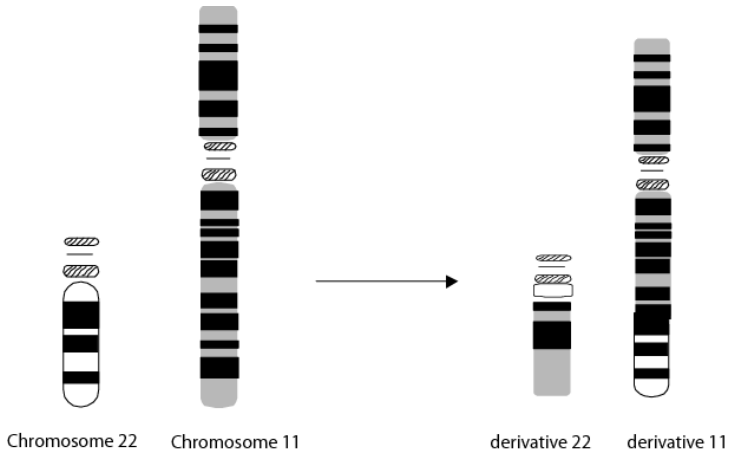


Figure 3. The constitutional translocation t(11;22)(q23;q11)

2.6 THE DIGEORGE /VELO-CARDIO-FACIAL SYNDROME (DGS/VCFS)

A group of developmental disorders, sometimes given the acronym of CATCH22 (C cardiac defect, A anomalous face, T thymus aplasia/hypoplasia, C cleft palate, H hypocalcemia and 22 for chromosome 22) and including DGS/VCFS, has been associated with microdeletions in 22q11.

DGS/VCFS is an autosomal dominant disorder with variable clinical expression that results from abnormal development of the third and fourth pharyngeal pouches during embryogenesis. Hemizygosity for a portion of 22q11 has been detected in 90% of VCFS/DGS patients. With an incidence of 1:4000 births (Burn, 1995), this is the most frequent deletion found in humans. Most of the deletions are *de novo*, it is estimated that only 5-10% of the deletions are inherited (Scambler, 2000).

The size of the deletion in 22q11 in most of the cases is 3Mb, the typically deleted region (TDR). Smaller deletions nested in the TDR are also common, and a few small deletions outside the TDR have also been reported (Rauch *et al.*, 1999; Saitta *et al.*, 1999). It is estimated that around 30 genes are located in this region and there are substantial efforts to elucidate the role of these genes in the etiology of DGS/VCFS (see Table 1).

Recently, one of the genes in the TDR, namely, the Tbx1 gene was targeted for inactivation in mice (Jerome & Papaioannou, 2001; Lindsay *et al.*, 2001; Merscher *et al.*, 2001). Heterozygous knockout mice displayed heart defects, which is the main characteristic of DGS/VCFS patients. Homozygous knockout mice died early and had a wide range of developmental anomalies encompassing most of the common DGS/VCFS features, including cleft palate, hypoplasia of thymus and parathyroid glands, conotruncal heart defects and abnormal facial structures. These results strongly suggest that the *TBX1* gene is involved in the etiology of DGS/VCFS, although it probably is not the only gene responsible for the wide spectrum of the syndrome. There are still a group of patients, which are not displaying the typical deletion who have been screened for mutations in the *TBX1* gene without success. A striking feature of DGS/VCFS is the extreme phenotypic variability that can be seen among patients carrying identical deletions. The presence of modifier loci has been discussed, and the *VEGF* gene has recently been reported to represent the first identified such modifier, affecting the risk of developing cardiovascular malformations in deletion patients (Stalmans *et al.*, 2003).

Table1. List of known genes in the DGS/VCFS typically deleted region

Gene	Description	Characteristics
PRODH	Proline dehydrogenase	Is involved in the degradation of proline. Animal models suggest that this gene may be involved in 22q11-associated psychiatric and behavioral phenotypes.
DGCR2	DGS critical region 2	Encodes a potential cell adhesion receptor protein
DGCR5	DGS critical region 5	Gene encoding a series of alternatively spliced transcripts expressed during human and murine embryogenesis, but with no obvious protein encoding potential. Disrupted by a translocation associated with DGS.
DGSI	DGS critical region gene	This gene is expressed in heart, brain, and skeletal muscle. Mutation analysis in deletion negative DGS/VCFS patients gave negative results.
DGCR6	DGS critical region 6	This gene detects a 1.1 kb transcript in many human tissues. The putative protein encoded by this gene shows homology with Drosophila gonadal protein (gdl) and with the gamma-1 chain of human laminin.
CLTD	Homo sapiens clathrin	Expressed in all fetal tissues tested and selectively expressed in certain adult tissues, particularly skeletal muscle.
SLC25A1	Solute carrier family 25, member 1	Also called citrate transport protein (CTP), responsible for the movement of citrate across the inner mitochondrial membrane. Expressed in many tissues, including liver, testis, ovary, and gut, but not in brain, skeletal muscle, or lung.
HIRA	histone cell cycle regulation defective, homolog A	Putative transcriptional regulator. The protein product contains WD40 domains, motifs thought to be involved in protein-protein interactions.
NLVCF	Nuclear localization signal deleted in DGS/VCFS	Encodes a deduced 206-amino acid protein containing 2 consensus sequences for nuclear localization signals and a putative leucine zipper near the N terminus. NLVCF is ubiquitously expressed in human adult tissues and in human and mouse embryonic tissues.
UFD1L	Ubiquitin fusion degradation 1-like	Encodes the human homolog of the yeast UFD1 protein, involved in the degradation of ubiquitin fusion proteins. This gene was identified as a downstream target of Hand2 transcription factor. Heterozygous Hand2 mice embryos have defects in heart development.
CDC45L	Cell division cycle homolog-like 45	In yeast Cdc45 protein is required for the initiation of DNA replication.
CLDN5	Claudin 5 (transmembrane protein)	The amino acid sequence has strong homology to the rat RVP.1 (rat ventral prostate) protein, a prostate-specific protein whose function is unknown. Computer analyses predicted 2 transmembrane domains. No introns.
GP1BB	Glycoprotein Ib (platelet), beta polypeptide	Platelet glycoprotein lb is a heterodimeric transmembrane protein. The role of GP1BB is unknown. The expression is most abundant in heart and brain.

TBX1	Brachyury variant B (TBX1), T-box 1	TBX1 is a transcription factor with a common DNA-binding domain, the T-box, involved in the regulation of developmental processes. Tbx1+/- mice have a high incidence of cardiac outflow tract anomalies. Tbx1-/- mice display a wide range of developmental anomalies seen in DGS/VCFS.
COMT	Catechol-O-methyltransferase	Catalyzes the transfer of a methyl group from S-adenosylmethionine to catecholamines. Plays a role in the metabolism of endogenous substances.
ARVCF	Armadillo repeat gene deleted in DGS/VCFS	The structure of ARVCF is closely related to the catenin family, which play important roles in the formation of adherend junction complexes. These data suggest that ARVCF is involved in protein-protein interactions at adherend junctions. ARVCF is ubiquitously expressed.
DGCR8	DGS critical region gene 8	Predicted, followed by confirmation of the exons by comparisons to protein, cDNA and EST databases.
ZNF74	Zink finger protein 74,	ZNF74 mRNA transcripts were detected in human and mouse embryos but not in adult tissues.
SERPIND1	Alternative names: Leuserpin 2, Heparin cofactor II.	Gene associated with thrombophilia due to heparin cofactor II deficiency.
CRKL	v-crk avian sarcoma virus CT10 oncogene homolog-like CRKL	Crkl -/- mice exhibited defects in multiple cranial and cardiac neural crest derivatives resembling the clinical manifestations of DGS/VCFS. Defects in CRKL-mediated signaling pathways are thought to be part of the molecular mechanism underlying DGS/VCFS.
LZTR1	Leucine-zipper-like transcriptional regulator, 1	Transcribed in several essential fetal organs. Several of its structural characteristics identified it as a transcriptional regulator.

3 AIMS OF THE PRESENT STUDY

This thesis addresses two related issues: Characterization of the constitutional translocation t(11;22)(q23;q11); and the use of array-CGH to detect small constitutional deletions.

The aims of this study were:

- To clone and characterize the breakpoints of the constitutional translocation t(11;22)(q23;q11).

- To investigate if breakpoint positions vary among carriers of the translocation.

- To find out if there are any gene(s) disrupted by the translocation in chromosomes 11 and/or 22, which could influence the development of breast cancer.

In the second part of this thesis, the aims were:

- To apply and further develop the array-CGH methodology for detection of gene dosage alterations on chromosome 22.

- To develop a second-generation high-resolution CGH-array covering the challenging 22q11 region. In particular, the aim was to compare the results obtained from use of genomic clone-based, cDNA-based, and repeat-free, non-redundant genomic PCR-based targets.

4 MATERIALS AND METHODS

4.1 PATIENT AND TUMOR MATERIAL

In the translocation t(11;22)(q23;q11) studies (papers I and III,) DNA from ten translocation carriers was used. The t(11;22)(q23;q11) was identified by standard cytogenetic procedures. All patients are females except patient 7 and 10, and all are unrelated except for patient 4 and 6, who are mother and daughter.

In the NF2 array-CGH study (paper IV) DNA samples from 116 patients were analysed. The patients were collected in collaboration with different laboratories in Bulgaria, Canada, Finland, France, Germany, Italy, Japan, Norway, Sweden, Switzerland, UK and USA. The patients were diagnosed according to the criteria from the NIH consensus conference in 1988 and are divided in mild, moderate and severe forms of NF2 (Evans *et al.*, 1992).

In the study using the full coverage chromosome 22 genomic microarray (paper V) different tumors and patient DNA were used to validate the performance of the array-CGH methodology. Two glioblastomas, fifteen acral melanomas and two dermatofibrosarcoma protuberans (DFSP) tumors were used to diagnose homozygous/heterozygous deletions and breakpoints of an unbalanced translocation. In addition, DNA from four DiGeorge patients and one NF2 female patient (p41), were also used.

In the DiGeorge study (paper VI), DNA from 10 patients was analyzed on the array. Three patients were referred to the Department of Clinical Genetics, Karolinska Hospital. The remaining seven DNA samples were prepared from commercially available cell cultures, one from Coriell Institute (Repository number GM07939) and six from The European Collection of Cell Cultures (ECACC), Accession numbers AC0461, AG0082, AC0001, AC0290, AG0124 and AC0004.

4.2 SCREENING OF GENOMIC LIBRARIES

To clone the translocation t(11;22) breakpoints we screened three genomic libraries:

- A PAC genomic library

- A cosmid library specific for chromosome 22
- A genomic library constructed from one of the translocation t(11;22)(q23;q11) carriers

The PAC library was constructed at the Roswell Park Cancer Institute, Buffalo, USA, (bacpac.med.buffalo.edu). It is a mixture of three libraries (RPC14, 15 and 16) in which the human genome is represented 14 times. It consists of 19 filters and each filter contains 18432 independent clones, the average size of the inserts is 116 kb.

The cosmid library specific for chromosome 22 was constructed in Lawrence Livermore National Laboratory (CA, USA), by flow cytometry. It is a mixture of two libraries, LL22NC01 and LL22NC03. The prefix "E" is used for the cosmids from the library LL22NC01 and "N" for the library LL22NC03. These libraries comprise 24 high-density nylon filters and contain approximately 19000 clones.

To obtain chimeric clones containing the regions of the breakpoints in chromosomes 22 and 11 a genomic cosmid library was constructed. High molecular weight DNA from one female carrier of the translocation t(11;22)(q23;q11) was used (paper III). The DNA was prepared using standard protocols and the vector used in construction of the library was SuperCos1 (Stratagene).

Different types of probes were used to screen these libraries:

- PCR amplified STSs from the public databases located in the vicinity of the translocation breakpoints in both chromosomes 22 and 11.

- Probes amplified by PCR from the end of the genomic clones. These "ends" were direct sequenced from the clones obtained from the library screenings or isolated by Δ-restriction cloning. Briefly, the restriction endonucleases *BamHI*, *XbaI* or *SacI* absent in the clone vectors pCYPAC2 and pPAC4 were chosen. PAC DNA was digested with one of these restriction enzymes and diluted to 100 ng/ml to avoid the re-ligation of the small fragments. Further a ligation reaction is performed with T4 ligase and transformed into XL Blue E. coli competent cells. Bacterial cells growth and PAC DNA containing the ends is isolated according to standard methods.

- For the screening of the translocation carrier genomic library, PCR amplified probes derived from the BAC clone 3009L19 (clone spanning the translocation breakpoints in chromosome 11) were chosen (paper III).

All the probes used for the screening of genomic libraries were radioactively labeled using random priming and Hot-PCR methods, according to standard protocols (Sambrook, 1989). Positive colonies were confirmed by colony PCR. DNA from BAC and PAC clones was extracted using the CsCl gradient method, Cosmid DNA was extracted with the Qiagen midi kit according to the protocol from the manufacturers (Qiagen, USA). Quality control and identification of identical clones was performed by digestion of DNA with *EcoRI*.
In the screening of the translocation carrier genomic library, all the positive clones, which were candidates for chimeric, translocation breakpoint-containing cosmids, were end-sequenced using the primers flanking the insert in SuperCos1 vector (sc1, GCA ATT AAC CCT CAC TAA AG and sc2, CCG CAT AAT ACG ACT CAC TAT). Two clones were identified using bioinformatic tools as chimeric cosmids, corresponding to the derivative 11 chromosome (cos4_der11) and derivative 22 chromosome (cos6_der22).

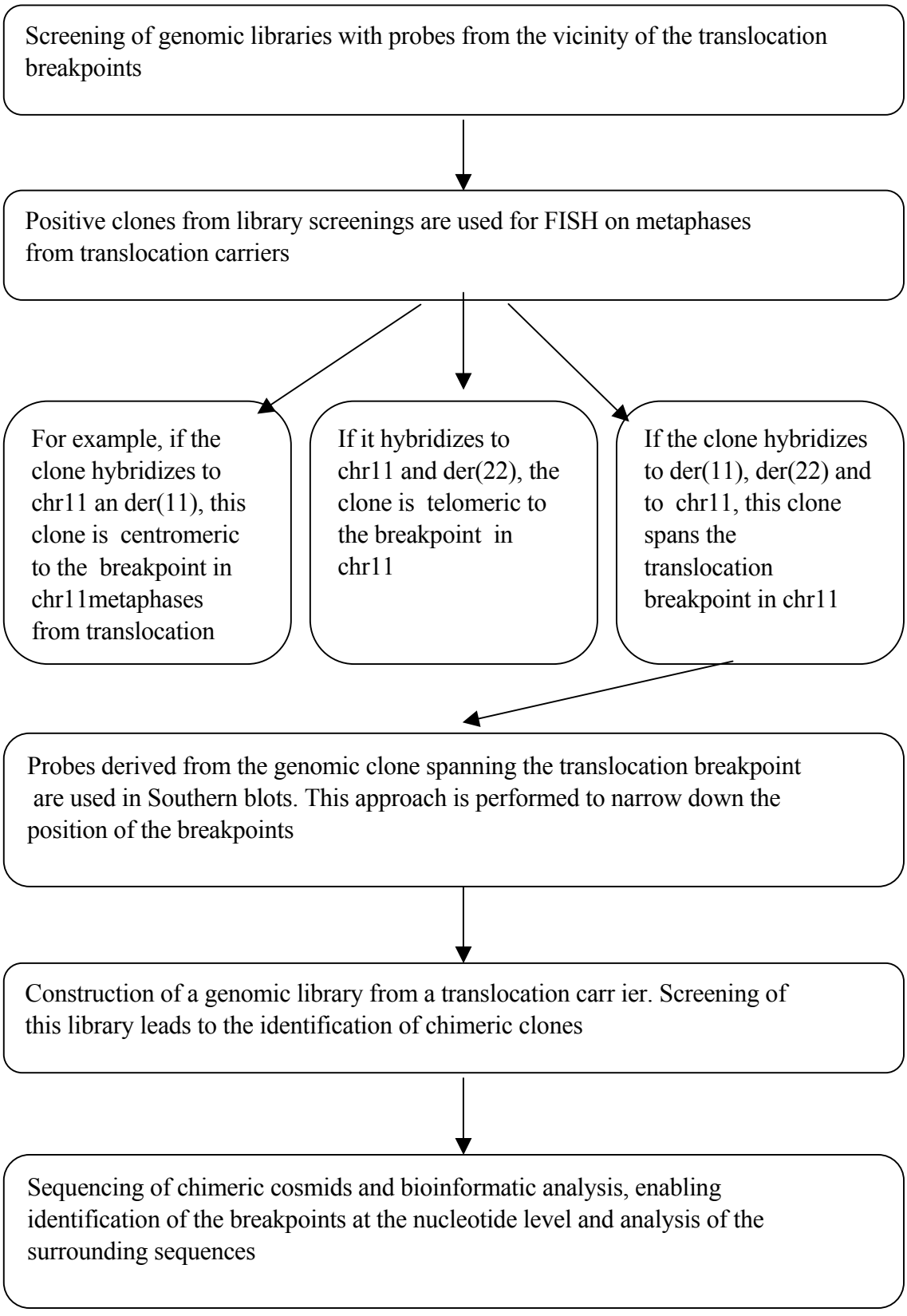
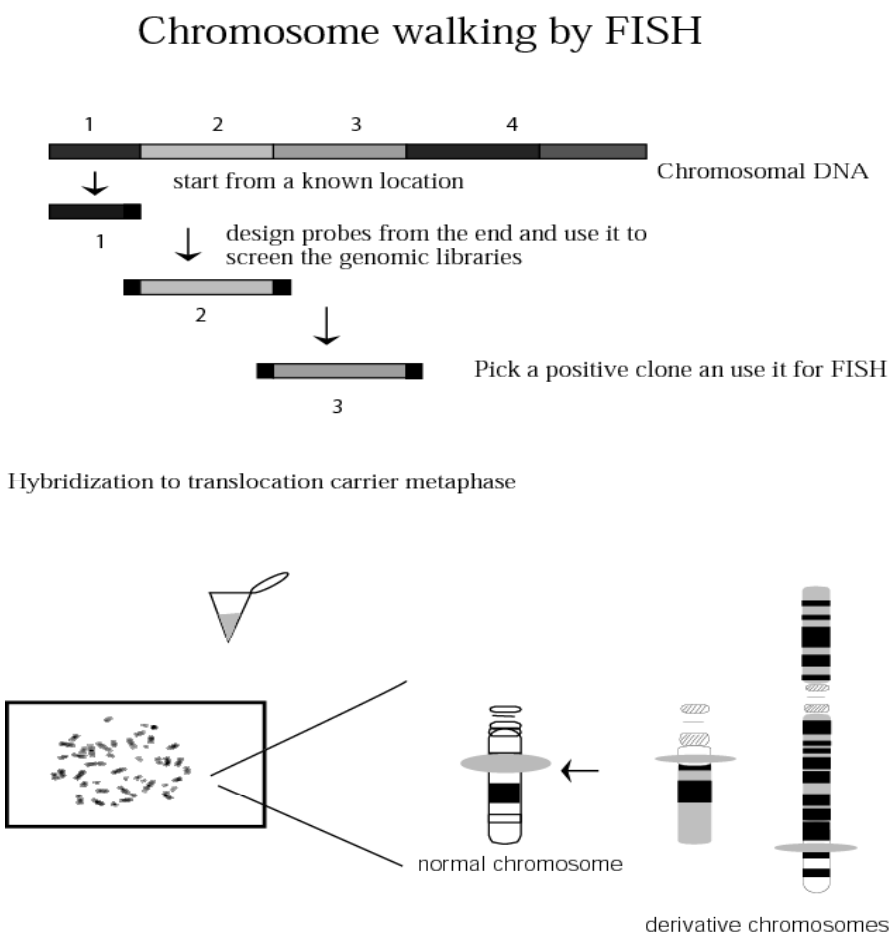


Figure 4. Strategy used for cloning of translocation breakpoints

4.3 CHROMOSOME WALKING BY FISH

Chromosome Walking is a widely used technique in physical mapping. The idea is to establish contigs composed of overlapping clones (so called tiling path) starting from points with known location. In order to map the region of the translocation breakpoints in chromosomes 22 and 11 we performed chromosome walking by FISH. For this purpose we screened genomic libraries as described above. Positive clones were used as probes in metaphase spreads from translocation carriers. For example, if using a chromosome 22 probe we obtain fluorescent signals from normal chromosome 22 and derivative chromosome 22, it means that the probe covers a region which is centromeric to the breakpoint in chromosome 22 or, if the probe gives signal in normal chromosome 22 and derivative chromosome 11 it means that the probe is telomeric to the breakpoint in chromosome 22, as it has moved to the derivative chromosome 11. If a probe spanning the translocation breakpoint is obtained, the signals in metaphase spreads from a translocation carrier will be in normal chromosome 22 and in both derivative chromosomes.



4.4 SOUTHERN BLOT ANALYSIS

We used Southern blots to localize probes near to the translocation t(11;22)(q23;q11) breakpoints in chromosome 22. Seven probes were used in the analysis, these were; c503, c754, c761, c99, c778, c73 and c439. The probes were amplified from BAC 2280L11 by standard PCR and radioactively labeled by random priming with α^{32}P (Feinberg & Vogelstein, 1984). Seven to ten micrograms of high molecular weight DNA from three control individuals and five carriers of t(11;22)(q23;q11) was digested with the restriction enzymes *PvuII*, *HindIII*, *BamHI* or *BglII*. The digested DNA was separated by 0.8% agarose gel electrophoresis and transferred to a nylon filter (Hybond-N+ Amersham). Hybridization was performed using standard conditions (Sambrook, 1989).

4.5 FLUORESCENCE *IN- SITU* HYBRIDIZATION (FISH)

FISH is a powerful technique that allows visualization of genetic alterations directly on nuclei/chromosomes. It is widely used in clinical diagnosis and research. It allows detection of imbalances by gains and/or losses of chromosomal segments, translocation breakpoints, and is commonly used in mapping. In this technique the target material is immobilized on a glass slide and the probe is labeled with fluorescent dyes, which are detectable by a fluorescent microscope. FISH has evolved considerably during the last decade and different targets can be used, such as; metaphase spreads, interphase nuclei and artificially extended DNA fibers. The resolution of FISH depends on the level of condensation and the nature of the target material, ranging from ~40 kb in metaphase spreads to 1-5 kb on the artificially extended chromatin fibers (paper III).

In this study (papers I and III), FISH was performed on metaphase chromosome spreads and artificially extended chromatin fibers from control individuals and carriers of the translocation t(11;22)(q23;q11). Cosmids, PACs and BAC clones obtained from the genomic library screenings were used in the FISH experiments. In the fiber-FISH, additional PCR products from a low copy repeat region of the PAC 52F6 were pooled and used as a single probe covering approximately 12 kb (paper III).

Metaphase FISH

Metaphase spreads were prepared from lymphocytes according to standard methods. The probes were blocked with Cot-1 DNA to suppress repeats. Hybridization was performed as previously described (Blennow & Tillberg, 1996). Probes were labeled by nick-translation

with modified nucleotides either with SpectrumOrange- or SpectrumRed-dUTP or Fluorescein-12-dUTP. The control probe for chromosome 22 was a telomeric cosmid N62A6 (Xie *et al.*, 1994) and the control for the chromosome 11 was a centromeric specific alpha satellite probe pHS53 (ATCC, MA, USA). For identification the chromosomes were counter-stained with DAPI (4,6 diamino-2-phenyl-indole) and the signal was visualized using a fluorescence microscope equipped with cooled CCD camera and the *SmartCapture* software package.

Fiber-FISH

Fiber-FISH was performed as described earlier (Fidlerova *et al.*, 1994, Yang, 1999 #54). Phytohemagglutinin-stimulated human lymphocytes were cultured for 72 h and harvested without colcemid treatment. Hypotonic treatment was performed in 0.075M KCl for 10 min at 37°C, and the cells were fixed in methanol: acetic acid (3:1). For releasing chromatin, fixed cells were spread on clean moist slides, and before evaporation of the fixative, slides were placed in PBS solution for 1 min. The slides were thereafter treated with 70% formamide in 2xSSC, pH 7.0, for 1 min, rinsed with 100% methanol, fixed with methanol: acetic acid (3:1), air-dried, passed through 70%, 95% and 100% ethanol, and air-dried again. Hybridization was performed as described above. Image analysis was performed using the programs Image-Pro Plus and Adobe Photoshop 4.0. Clones used for fiber-FISH were labeled either with digoxigenin and detected by using FITC-conjugated anti-digoxigenin antibodies (green) or with biotin and detected by rhodamin-conjugated avidin (red).

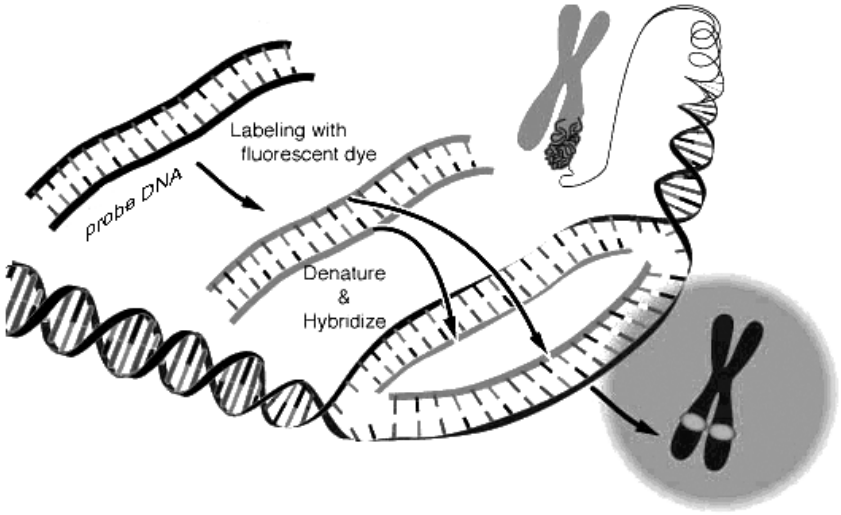


Figure 6. Principle of fluorescence *in-situ* hybridization (FISH). The probe is labeled with fluorochromes, denatured and hybridized to a target molecule. The target molecule can be a chromosome in metaphase (shown in this figure example), interphase or DNA fibers.

4.6 AUTOMATED SEQUENCING

In this study we used two kinds of sequencing approaches:

1. Large-scale genomic sequencing by shotgun cloning. In this method the goal is to create subclones with small inserts from large genomic clones, which have inserts larger than the average read length of an individual sequencing reaction. DNA is randomly sheared in small pieces by nebulization/sonication and subcloned into PUC18 vectors. DNA is extracted from the subclones and used as template for the sequencing reaction. The sequences are assembled and the entire sequence of the original large genomic clone is generated.

2. End-sequencing of genomic clones. Prior to the entire clone sequencing, we first generated end-sequences of the large clones. For this purpose larger amounts of template (500 ng-1 ug) and an increased number of cycles (50-100) in the sequencing reaction are required.

The sequencing technique widely used today is mainly derived from the enzymatic method of sequencing, known as the Sanger-Coulson method (Sanger *et al.*, 1977). The DNA template to be sequenced is amplified by PCR, the reaction is performed with the presence of dideoxynucleotides (ddNTP) acting as chain-terminating inhibitors in a cycle sequencing reaction. The ddNTPs contain a hydrogen molecule in the 3´ carbon of the deoxyribose, which prevents the attachment of additional nucleotides. The DNA polymerase enzyme randomly adds both deoxynucleotides and dideoxynucleotides that are complementary to the DNA template. The "new" strand is terminated by the addition of a ddNTP. The result is many copies of every possible fragment each terminated by a ddNTP. The dideoxynucleotides are labelled with different fluorochromes and the sequencing reaction is separated on 0.2 mm Long Ranger acrylamid gel. Each ddNTP emits light of certain wavelength when excited by a laser beam. This light is recorded and shown as a simulated gel image. A computer program interprets the raw data and the output is a chromatogram in which every peak will represent a nucleotide.

4.7 BIOINFORMATICS AND DATABASES

Numerous programs were used for assembly and analysis of sequences generated in this work. Individual sequences were assembled into contiguous sequences using the *asp* and *pregap4/gap4* programs from the Staden package (Staden, 1994) on a UNIX computer. When screening the genomic libraries (papers I-III) STSs were used, these were selected according to their genomic location from the Whitehead institute (http://www-genome.wi.mit.edu/cgi-bin/contig/phys_map), the genome database (http://www.gdb.org/) and MapViewer at the NCBI (http://www.ncbi.nlm.nih.gov/). Sequences of the above markers were used to pick large genomic segments of the *nr* (Non-Redundant) and *htgs* (High Throughput Genomic Sequences) databases using the BLAST family of programs at the NCBI/NIH server (www.ncbi.nlm.gov/BLAST). Repetitive sequences were filtered out using the program *RepeatMasker* (repeatmasker.genome.washington.edu). Primers were designed using the program Oligo 5 (Molecular biology Insights, CO, USA). For sequence comparison Dot-plot was performed with the program *dotter* (Sonnhammer & Durbin, 1995). Gene prediction was performed using the *GeneScan* program (genius.embnet.dkfz-heidelberg.de). The formation of secondary structures with sequences from the chimeric cosmids was predicted using the *mfold* program (http://BiBiServ.TechFak.Uni-Bielefeld.DE/mfold/) (paper III).

4.8 COMPARATIVE GENOMIC HYBRIDIZATION TO MICROARRAYS (MICROARRAY- CGH)

The conventional method of comparative genomic hybridisation (metaphase-CGH) is hybridized to metaphase spreads (Kallioniemi *et al.*, 1992; Kallioniemi *et al.*, 1993). Today this is a well-established method for detecting large chromosomal rearrangements. However, due to the high level of condensation and supercoiled structure of DNA in the chromosomes, the resolution in detecting gene copy number changes is no less than 5-10 Mb for deletions and approximately 2 Mb for copy number amplifications/gains.

The principle of this technique is a competitive hybridisation between test DNA and control DNA, labeled with different fluorochromes (e.g. Cy3 and Cy5) in the presence of unlabeled Cot1 DNA to block the repeat sequences present in the spotted target DNA on a glass slide. Thus, if we have an heterozygous or homozygous deletion i.e. the number of copies of a particular target locus has decreased (one or zero compared to two), a small amount of test DNA will hybridize to those spots on the microarray that represent the locus involved in the deletion, whereas comparatively larger amounts of control DNA will hybridize to those same spots. The fluorescence ratio of the test and control DNA hybridized to the target indicate the relative copy number of the test DNA at a specific location.

The development to microarray-based form of CGH is a dramatic improvement of the method. Already in the first publications describing the use of microarray-CGH, the resolution reached the same level as for FISH-analysis (Pinkel *et al.*, 1998; Solinas-Toldo *et al.*, 1997), while the throughput realistically enables genome-wide analysis on a single slide. In microarray-CGH pieces of DNA serve as targets on a solid surface. The targets are printed as individual spots and each spot has a predefined chromosomal location. In this work we have used three main types of targets spotted on the glass slide:

1. Genomic clone-based array targets, such as BACs, PACs, cosmids and fosmids.
2. Non-redundant, repeat free pools of PCR-amplified genomic DNA.
3. cDNA-based targets, such as single ESTs spotted on the array and pools of amplified inserts from EST-clones.

As mentioned above, microarray-CGH has dramatically increased the resolution in detecting gene copy number variation, compared to the conventional CGH methodology, which uses

metaphase spreads. When genomic clones are used it is possible to detect deletions as small as the size of a cosmid insert, which is approximately 40 kb. We show that the resolution can be increased even further using non-redundant, repeat-free PCR amplified pools (papers VI and V)

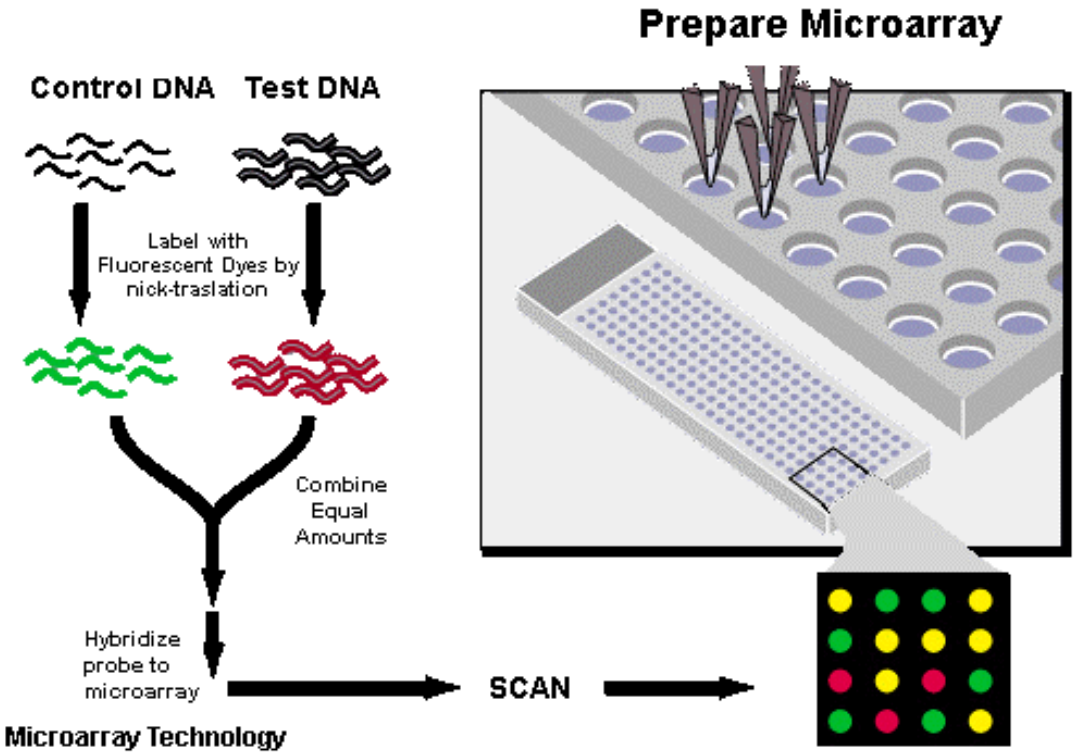


Figure7. Comparative Genomic Hybridization to microarrays

5 RESULTS AND DISCUSSION

5.1 MAPPING AND CLONING OF THE CONSTITUTIONAL TRANSLOCATION t(11;22)(q23; q11) (PAPERS I AND III)

It has been reported that in a group of t(11;22)(q23;q11) translocation carriers there is an increased predisposition for development of breast cancer (Lindblom *et al.*, 1994). This fact has been the starting point of our interest in the characterization of this translocation. If a gene or genes are involved in a translocation event, the result may be a juxtaposition of two genes, which were not previously located in the vicinity of each other. This may lead to the activation of a specific gene or the translocation may disrupt a gene altering its function. Several examples of such events are known to initiate tumorigenesis. In an attempt to clone the t(11;22)(q23;q11) breakpoints, we performed chromosome walking by FISH in both chromosomes 22 and 11. Starting from the better-characterized chromosome 22, we screened the genomic libraries and the electronic databases several times until we reached a point in which all the clones obtained mapped to the same location. Mapping of this translocation has been a very arduous task, mainly due to the presence of repetitive sequences in the region where the breakpoints occur.

5.1.1 Mapping of the translocation breakpoints in 22q11

In chromosome 22 we mapped the translocation to a region between markers D22S134 and D22S264. We obtained two BAC clones, 2280L11 and 41C4, that both produced positive signals by FISH analysis to metaphases from translocation carriers, they hybridized to normal 22 and to both derivative chromosomes. Thus, we assumed that the two BAC clones were spanning the translocation breakpoints in chromosome 22. Later, interphase FISH with the same BAC clones presented numerous signals showing that these clones contain LCRs and hybridize not only to the region where the break occurs in chromosome 22, but they also hybridize to other LCRs, proximal and distal to the t(11;22) translocation breakpoint. We concluded that these BAC clones contain sequences similar/identical to the sequences involved in the translocation breakpoints in chromosome 22. Surprisingly, when we sequenced both BAC clones 2280L11 (AC011718) and 41C4 (AC012331), analysis of the repeat-masked sequences revealed that although they share stretches of high score matches (above 99%) and hybridize to the same places in chromosome 22, they do not fully overlap. They probably map to the same gap in the sequence of chromosome 22 but despite all the efforts in characterizing

this region in 22q11 these BAC clones have not yet been exactly mapped. One additional confirmation that sequences similar/identical to the sequences contained in the two BAC clones mediate the translocation t(11;22) is the fact that probes derived from the BAC 2280L11 were used in Southern blot analysis, where we could detect differences between carriers of the translocation and normal controls (paper I).

5.1.2 Mapping of the translocation breakpoints in 11q23

On chromosome 11, we mapped the translocation to a region between markers D11S1340 and WI-8564. Walking on this chromosome was much easier and faster, leading to the isolation of the BAC clone 3009a19. Metaphase FISH analysis in translocation carriers with this clone resulted in signals on normal chromosome 11 and in both derivative chromosomes, i.e., this clone spans the translocation breakpoint in chromosome 11 in the nine patients analyzed (Table 1 in paper I). Interphase FISH revealed single signals in non-translocation controls confirming that this BAC clone in fact spans the translocation breakpoint in chromosome 11.

We also used the BAC 3009a19 for fiber-FISH studies and detected deletions in this BAC clone. In all analyzed images a significantly weaker signal was detected in the central part of this clone (Figure5b in paper III). The likely explanation is that this DNA used for fiber-FISH experiments is a mixture of BACs with full insert and a substantial proportion of BAC molecules containing deletions, caused by the instability of this region upon propagation in bacteria. This is perhaps due to the presence of AT-rich palindromes sequences in this locus (see below). The same phenomenon has been reported in the literature for other BACs cloned from the same region of 11q23 (Edelmann *et al.*, 2000; Kurahashi *et al.*, 2000b).

5.1.3 The breakpoints of the constitutional translocation t(11;22)(q23;q11) occur at AT-rich palindromic sequences

The identification of the borders of chromosome 22 and 11 was not straightforward as this region is AT-rich and highly unstable upon culturing of genomic clones in bacterias. In fact, all the genomic clones derived from normal chromosome 11, which are deposited in the databases, contain small deletions. In order to clone the translocation t(11;22) breakpoints at the nucleotide level we constructed a genomic library from one translocation carrier. This library contained clones from the region of juxtaposition in the derivative chromosomes. To screen this genomic library we used probes derived from sequences from BAC 3009a19 known to be in the vicinity

of the breakpoint in chromosome 11. As a result, we isolated two chimeric cosmids representing the region of the juxtaposition in both derivative chromosomes: cos4_der11 and cos6_der22. Both chimeric cosmids were fully sequenced and contain 42 and 40 kb, respectively. In both chimeric cosmids, the portion of chromosome 11 was substantially bigger than the part derived from chromosome 22. This is probably another indication of the difficulty in cloning of this region from chromosome 22. The region where the break occurs in 22q11 is located within a gap in sequence of 22q11. Therefore, in order to identify the sequences derived from chromosome 22 and 11, we aligned both chimeric cosmids with normal chromosome 11 sequence. This allowed us to determine the exact position of translocation between chromosome 22 and 11 in the derivative chromosomes.

One important achievement in the isolation of the chimeric cosmids is that we generated additional ~14 kb of sequence from chromosome 22, which is a part of the unclonable gap. Computer analysis of the chromosome 22 and chromosome 11 sequences derived from the chimeric cosmids revealed the formation of hairpin-like structures or palindromes. In the case of chromosome 22, the palindrome was very strong containing 852 bases with only 8 mismatches. The palindrome formed from the sequences of chromosome 11 is 166 bases with two mismatches. After extensive analysis of sequences from these palindromes we conclude that the breakpoints are at the tip of the palindromes in both chromosomes (Figure 3, paper III). This suggests that the centers of palindromes are susceptible to breaks and are involved in the mechanism of the formation of this translocation. It is known that the presence of inverted repeats or palindromes cause genetic instability. The degree of this instability conferred by the palindromes is proportional to the length of the repeat, the distance between the repeats and the homology shared by them as it has been shown in bacteria (Leach, 1994), yeast (Gordenin *et al.*, 1993; Lobachev *et al.*, 1998) and mouse (Zhou *et al.*, 2001), but little is known about palindromes in humans. Recently the breakpoints of another constitutional translocation, the t(17;22) in a patient with NF1 have been cloned. In the latter translocation, the breakpoint in chromosome 22 is at the same location as for the constitutional translocation t(11;22). The chromosome 17 breakpoint is located within intron 31 of the *NF1* gene and also contains a palindromic structure (Kurahashi *et al.*, 2003). This suggests that there is a common mechanisms responsible for the etiology of both constitutional translocations t(11;22) and t(17;22), which is mediated by formation of inverted repeats/palindromes.

Furthermore, fiber-FISH analysis of four translocation carriers and a normal control was performed to determine whether the translocation breakpoints occur in the same region. In this

analysis we used the BAC 3009a19 and both above described derivative cosmids as probes. The derivative cosmids are originally cloned from the translocation carrier 6 (paper I) and the translocation carriers used in this analysis are four additional unrelated subjects (cases 2, 3, 5 and 10, paper I). The probes were used in pairwise combinations and labeled with green and red fluorochromes. In this study, we could not detect any signal derived from the derivative 11 cosmid cos4_der11, which overlaps with the signal from BAC 3009a19 on the DNA fiber originating from derivative 22 chromosome (Figure 5c, panel 2, paper III). Similarly, there was no signal derived from the derivative 22 cosmid cos6_der22, which overlaps with the signal from BAC 3009a19 on the DNA fiber originating from derivative 11 chromosome (Figure 5c, panel 4, paper III). Therefore, we concluded that the breakpoints are in the same location all patients analyzed, at the resolution of fiber FISH. In our hands, the resolution of fiber FISH experiments is in the range of 2 kb. This conclusion is derived from hybridization experiments using the cosmid cos6_der22 to the normal chromosome 22, which contains only ~2 kb of chromosome 22 sequences (Figure 5a, panel 2, paper III).

As mentioned previously, our initial aim in the cloning of the constitutional translocation t(11;22)(q23;q11) was the identification of a gene(s), which might be disrupted by this translocation. Such a gene could be involved in the etiology of breast cancer in some translocation carriers. LOH studies in sporadic breast tumors have pointed to 22q11 and 11q23 as potential sites harboring breast cancer-related genes (Laake *et al.*, 1999; Larsson *et al.*, 1990). Our analysis reveled that no gene is disrupted either in chromosome 22, or in chromosome 11. However, this does not eliminate the possibility that genes located in the vicinity of t(11;22) translocation breakpoints are involved in the etiology of breast cancer, by e.g. positional effect disturbing expression/regulation of neighboring gene(s). Thus, the question regarding possible predisposition to breast cancer in this group of translocation carriers remains open.

5.2 SEQUENCING OF CHROMOSOME 22 (PAPER II)

Chromosome 22 is the first chromosome that was sequenced. This report was published in December 1999, as a result of a substantial international effort in which 218 investigators from laboratories in UK, USA, Japan, Canada and Sweden were involved. We reported the sequence of 12 contiguous segments spanning 33.4 Mb. There are 11 gaps in the sequence of 22q reported in 1999, none of these gaps is considered to be bigger than 150 kb. Several groups are intensively working with aim of closing these gaps. However, despite this, only one gap in the

Cat Eye syndrome region from 22q11 has been successfully closed to date by The Advanced Center for Genome Technology, University of Oklahoma (USA).

The sequencing of chromosome 22 was performed using the clone-by-clone strategy and the shotgun sequencing methodology (see material and methods). Analysis of sequences of 22q revealed that the human chromosome 22 contains a total of 679 genes, of which 247 are known genes, 150 are related genes, 148 are predicted genes and 134 are pseudogenes. Our laboratory has worked in several regions of chromosome 22, including the regions commonly deleted in meningioma, covering the neurofibromatosis 2 gene, the region involved in the aetiology of dermatofibrosarcoma protuberans, the 22q11 region which is often deleted in DGS/VCFS syndrome and also involved in the generation of the constitutional translocation t(11;22). Some of the clones derived from screening of genomic libraries performed during mapping of the constitutional translocation t(11;22)(q23;q11) are a part of the minimal tiling path of chromosome 22. These include the PAC 158L19 (AC006547), PAC 52f6 (AC005500) and PAC 888c9 (AC005663). Other clones are located in the chromosome 22, but their exact location has not yet been established. These are: BAC 2280L11 (AC011718), BAC41c4 (AC12331), BAC 2262a2 (AC012330) and the cosmid c31p7 (AC012398).

As mentioned earlier, 22q11 is one of the most difficult regions to sequence on 22q, due to presence of LCRs, which made the chromosome walking very difficult. An additional difficulty, specifically related to the t(11;22) translocation breakpoint region, is presence of unclonable and unstable sequences. We screened the genomic libraries several times, but we have been unable to cover this region with contiguous clones. The clones from this region stretching towards the gap share 98-99% sequence identity with each other and with additional 22q segments containing LCRs. This makes the mapping process very difficult as hybridisation experiments, amplifications by PCR and screening of genomic libraries are of little use when the level of sequence similarity is so high. A combination of different novel methodologies will be necessary to apply, for a successful closure of the remaining gaps in the human genome.

Availability of the entire sequence of chromosome 22 allowed chromosome-wide studies of specific duplications in this chromosome. A recent study estimate that 10.8% of 22q sequence is duplicated (defined as 3.7 of 33.8 Mb, larger than 1 kb and with sequence identity of at least 90%) (Bailey *et al.*, 2002). Segmental duplications vary in size from one to a few hundred kb; their role in the whole genome is seen as a major pathway of genome evolution. The structure of the duplications is often very complex because of the large number of transpositions,

juxtapositions and other rearrangements. Unlike common repeats (such as Alu, LINE, etc.), LCRs are usually larger in size (between 10-400 kb). It has been estimated that approximately 5% of the sequences in the human genome are LCRs (Stankiewicz & Lupski, 2002). In the case of LCRs from chromosome 22, we believe that these structures are the basis of instability in the 22q11 region. During meiosis the regions containing LCRs may undergo homologous recombination with highly paralogous segments producing rearrangements in the gametes, such as deletions or duplications.

5.3 THE USE OF MICROARRAY-CGH FOR DETECTION OF SMALL CONSTITUTIONAL DELETIONS (PAPERS IV AND V)

An increasing number of genes causing human diseases are known to be involved in micro-deletions/micro-gains of chromosomal segments (less than a few Mb in size), resulting in gene copy number changes. Until recently, no method has been available to comprehensively screen for these types of rearrangements. Today, the method of comparative genomic hybridization to microarrays (array-CGH) opens up new possibilities for high-resolution and large-scale identification of gene dosage imbalances. In the case of cancer, aberrations at the gene copy number level play an important role in the initiation and progression of the disease. Characterization of the deletions and/or gains a cancer cell contains may help in the identification of tumor suppressor and/or oncogenes, which are lost or amplified, respectively.

In papers IV and V, we demonstrate the usefulness of the array-CGH methodology in various clinical and research-related applications. Paper IV is focused on the use array-CGH in detection of deletions in a large series of neurofibromatosis type 2 (NF2) patients. NF2 is an autosomal dominant disorder, affecting approximately 1 in 40000 individuals. The *NF2* gene was identified in 1993 (Rouleau *et al.*, 1993). It was subsequently shown that bi-allelic mutations in this gene are responsible for the etiology of NF2. Since then, several mutation studies have been performed on both familial and sporadic NF2 cases. It is estimated that the mutation frequency in NF2 patients is up to 85% (Zucman-Rossi *et al.*, 1998). Approximately 25% of the NF2 patients have constitutional deletions that (partly or entirely) remove one copy of the *NF2* gene. The patients used in this study were classified into three groups: severely affected, moderately affected and mildly affected patients according to the diagnostic criteria suggested in 1992 (Evans *et al.*, 1992). By these criteria the severely affected patients (Wishart) have an early onset, rapid course of disease and develop bilateral vestibular schwannomas as well as multiple other tumors. The mildly affected patients (Gardner) have a

late onset, more benign course and usually only bilateral vestibular schwannomas. Patients who do not clearly fall into either category are considered as moderately affected.

A microarray covering a 7 Mb interval from chromosome 22 in the vicinity of the *NF2* gene was constructed. We spotted 104 genomic clones from chromosome 22 and seven control clones from chromosomes 3, 5 and X on the array. We performed the following validation experiments, in order to assure that the array is capable to quantitatively detect gene copy number variation:

1. Hybridization of DNA from a normal male (XY) against normal female DNA (XX). This shows the values for one copy on the X chromosome and two copies for all other chromosomes.
2. Hybridization of DNA from a patient (JP) with a known constitutional deletion of approximately 7 Mb (Bruder *et al.*, 1999).
3. Hybridization of DNA from a meningioma tumor (Tumor 24), which contains both homozygous and heterozygous deletions in chromosome 22.

DNA from 116 patients was hybridized to the array, of these 83 were severely affected, 16 were moderately affected and 17 were mildly affected. We detected deletions in 24 patients, eight severe, nine moderate and seven mild cases. The smallest deletions had the size of a cosmid insert, i.e., approximately 40 kb. These small deletions were detected in patients with both severely and mildly affected phenotypes. In patients having moderate phenotypes the deletions varied from 120 kb to 6.6 Mb. Correlation between the type of the *NF2* gene mutation and the patient phenotype has previously been postulated (Ruttledge *et al.*, 1996). Our results do not support the hypothesis that there is a correlation between the type of mutation affecting the *NF2* gene and the disease phenotype.

In paper V, we constructed the first comprehensive, full-coverage and high-resolution chromosomal genomic microarray. This chromosome 22 array covers approximately 35 Mb, representing 1.1% of the genome and 2.4% of known genes, with an exceptional average resolution of 75 kb. To demonstrate the utility of the array, we have applied it to profile acral melanoma, DiGeorge syndrome and neurofibromatosis 2. We accurately diagnosed homozygous/heterozygous deletions, amplifications/gains and *IGLV/IGLC* locus instability. We further identified using this array the 14-3-3 eta isoform as a candidate tumor suppressor in glioblastoma. Two glioblastomas were hybridized and we could detect large heterozygous deletions in both cases. In one of these glioblastomas, we also detected a homozygous deletion

in one cosmid clone. This cosmid contains the 14-3-3 protein eta isoform gene, which is a member of a protein family known to play role in mitosis, apoptosis and various signal transduction pathways. This gene has not earlier been associated with glioblastomas. Thus, it is an important finding as it demonstrates the usefulness of our array-CGH in the detection of novel candidates for tumor suppressor genes. We also applied this array to profile dermatofibrosarcoma protuberans (DFSP). At the molecular level DFSP is identified by the presence of the *COL1A1*/*PDGFB* fusion gene, which is formed by a reciprocal translocation t(17,22)(q22;q13) or supernumerary ring chromosomes containing material from chromosomes 17 and 22. We hybridized two DFSP tumors and detected the breakpoints of the t(17,22)(q22;q13) to one cosmid clone containing the *PDGFB* gene. This cosmid and all the following clones to the telomeric side display the fluorescent ratios consistent with a single gene copy level. This is a very quick way to detect an imbalanced translocation and could potentially be used for diagnostic purposes of DFSP. A broad spectrum of issues can be approached using this array. Chromosome 22 is a well annotated and gene-rich autosome and contains a considerable number of uncharacterized disease genes, e.g. familial schizophrenia susceptibility, glioblastoma, ependymoma (and other types of glioma), meningioma, schwannomatosis, pheochromocytoma, Wilm's tumor, breast and colon cancer (Dumanski, 1996). It is therefore important to associate genes from this chromosome to specific 22q-related conditions and this array will be instrumental towards this goal. Furthermore, comprehensive epigenetic profiling (changes affecting metylation status of genes) of 22q-located genes and high-resolution analysis of replication timing across the entire chromosome can be performed using our array.

Two methodological advances in array construction were also developed and validated. One of the main constraints of array-CGH protocols is a low yield of DNA obtained from genomic clones, which often requires costly repetitions of large-scale preparations. The goal was to test the performance of DNA amplified with phi-29 polymerase, which is a highly processive, proof-reading and strand displacing DNA polymerase derived from the phi-29 phage of *Bacillus subtilis*. Large linear and circular DNA molecules can be efficiently amplified using the rolling circle amplification principle (Dean *et al.*, 2002; Dean *et al.*, 2001). The results of phi29 amplified DNA from all so far performed experiments suggest that phi29 polymerase-based protocol performs well and is a reliable way for array construction, which entails a considerable advantage with regard to work input.

The other successful development was the application of a strictly sequence defined, repeat-free, and non-redundant strategy for array preparation. Chromosome 22 is a suitable experimental platform for developing new technologies in array-CGH. It is a challenging subject as the tiling path of 22q is rich in inter- and intra-chromosomal duplications (also denoted as Low Copy Repeats), many regions with a very high content of common repeats, pseudogenes and numerous genes which have closely related paralogs in other parts of the genome. This restricts the analysis of this important chromosome. Furthermore, current protocols for array-CGH rely on human Cot-1 DNA for suppression of background derived from common repeats. However, Cot-1 DNA is expensive and the quality of different commercially available batches is extremely variable. Thus, a decrease in the amount of Cot-1 DNA, or its full elimination from array-CGH is desirable. Additional drawbacks associated with genomic clones include their unavailability, contamination with bacteriophages, inserts that are prone to rearrangements, and purchasing of expensive clones from private distributors. The rationale was to test an approach, which would overcome above limitations of array-CGH. Our repeat-free strategy eliminates a majority of the above limitations, allows also a dramatic increase in array resolution (currently down to a 10-20 kb genomic segment) and analysis of any locus; disregarding common repeats, genomic clone availability and sequence redundancy. Results from our experiments confirm that the pools of repeat-free and non redundant PCR-fragments perform as reliable replacements for genomic clones.

5.4 CONSTRUCTION OF AN IMPROVED ARRAY COVERING THE CHALLENGING 22q11 REGION, FOR DETECTION OF MICRODELETIONS IN DGS/VCFS PATIENTS (PAPER VI)

To date, two distinct approaches used in microarray CGH: genomic DNA-based and cDNA-based platforms. The genomic approach is already well established and several studies are published using this platform (Buckley *et al.*, 2002; Geschwind *et al.*, 1998; Hui *et al.*, 2001; Pinkel *et al.*, 1998; Solinas-Toldo *et al.*, 1997). In the case of cDNA-based array-CGH, only a few publications have shown successful application of this platform (Beheshti, 2003; Heiskanen *et al.*, 2000; Kauraniemi *et al.*, 2001)**,** since the first article was published in 1999 (Pollack *et al.*, 1999). The main disadvantage of this approach are difficulties in obtaining quantitative data for gene copy number variation and a high rate of false positive/negative results (Beheshti, 2003).

As mentioned above, the 22q11 chromosomal segment containing the region that is typically deleted in DGS/VCFS, is challenging as it contains low copy repeats, segmental duplications, unclonable sequences and has a high content of common repeats. Not surprisingly, the typically deleted region (TDR) proved to be particularly difficult to analyze using our above described whole chromosome 22 array (paper V). We attempted to design an improved array, specifically tailored to detect gene-dosage imbalances of the TDR in the 22q11 segment. In this process, we compared the use of genomic clones, repeat-free pools of amplified genomic DNA, and cDNA clones (single and in pools) containing ESTs derived from genes in this region, as targets on the array.

We constructed a 6 Mb microarray including the 3 Mb from the TDR in DGS/VCFS. On this array, genomic as well as cDNA targets were printed (Table 2). The genomic targets were genomic clones such as BACs, PACs, cosmids and fosmids. In addition, PCR pools from three genomic loci were constructed, containing repeat-free and non-redundant genomic sequences amplified by PCR. The cDNA targets were single cDNA as well as pools of cDNAs. We pooled between 2 and 11 different cDNA clones to form individual spots on the array (Table 2).

Table 2. List of printed targets on the 22q11 array

Number of targets	Type of targets
49	Genomic clones from DiGeorge region in 22q11
9	Genomic control clones from X chromosome
9	Genomic control clones from other loci in 22q
3	Genomic control clones from other autosomes
21	Single cDNA clones from DiGeorge region in 22q11
4	Single cDNA clones from other 22q loci
14	Single cDNA clones from X chromosome
24	Single cDNA clones from other autosomes
9	cDNA pool of DiGeorge syndrome region
2	cDNA control pool from chromosome X
6	cDNA control pool from other autosomal chromosome
3	Pools of PCR products derived from *ARVCF*, *TBX* and *SERPIND1* genes, within TDR in 22q11

To validate the performance of this array, in detecting gene copy number changes, we co-hybridized DNA from a normal male (XY) and a normal female (XX). For all the genomic clones on the array the average normalized inter-locus fluorescence ratio (ANILFR) was as

expected (two copies: ANILFR 1.01±0.05; one copy: ANILFR 0.67±0.05). For the cDNA pools and for the repeat-free, non-redundant genomic PCR pools the ratios were ANILFR 1.01±0.05 for two copies, and ANILFR of 0.67±0.05 for one copy. Single cDNAs and pools of only two cDNAs did not display reliable results in quantitative detection of gene copy number alterations.

The three PCR amplified genomic pools, covering repeat-free genomic regions of 11.48, 12.96 and 15.45 kb, represent the *SERPIND1*, *ARVCF* and *TBX1* genes, respectively. These three genes are located within the TDR of DGS/VCFS. All three PCR pools could detect heterozygous deletions (one gene copy). We demonstrated that as little as 11.5 kb repeat free pooled PCR products can be used in reliable detection of gene copy number variation using the array-CGH methodology. It is estimated that the human genome contains around 50% repeated sequences. By extrapolation, this indicate that we could perform analysis of the human genome with an average resolution of 1 measurement point for every 25 kb using this approach.

The use of cDNAs as target in array-CGH would allow analysis of both expression and gene copy number in parallel. Such a combined platform is still under development. We could not obtain reliable, quantitative results from spots on the array containing single ESTs. Recently, another group also reported difficulties in detecting low copy number imbalances using single cDNAs. Their analysis showed that gene copy number variations of less than 10 to 20 copies could not be reliable detected (Beheshti, 2003). To test the minimum amount of sequences necessary for a reliable detection of haploid/diploid gene dosage level, we pooled cDNAs. The combined length of sequence in each pool varied from 2.5 kb (pool of two cDNAs) to 19.1 kb (pool of eleven cDNAs) (Table 2, paper VI). Only when we used a minimum of 3.5 kb of EST sequence (three clones pooled), we could quantitatively detect heterozygous deletions. We hybridized to the array DNA derived from lymphoblastoid cell lines of 10 DGS/VCFS patients. We observed that in six cases the deletion encompassed 3 Mb, which is the entire TDR. In two cases the deletions appeared to be very small and were detected only by one or two genomic clones. In two additional cases, there were no clones displaying fluorescent ratios consistent with deletions (Figure 1, paper VI).

In our DGS/VCFS array, 34 out of 49 genomic clones cover TDR. In the case of the six patients displaying the 3 Mb typical deletions, 16 clones were always deleted, two always failed to detect the deletion and 16 produced variable results. We believe that there are several reasons for these inconsistent results. For the majority of the clones that were unreliable in predicting deletions, there is a plausible explanation as to why the fluorescence ratios deviated from the

expected values for hemizygosity. For instance, PAC p52f6 contains LCRs. Furthermore, a majority of the remaining clones, which show deviant fluorescence ratios, are cosmids/fosmids with relatively short inserts and high content of common repeats. Our previous analysis of the entire chromosome 22 on the genomic array suggested that cosmids, with a high content of common repeats, are more prone to deviations, as compared to large insert genomic clones. However, in one instance, bioinformatic analysis of the genomic sequence from PAC p158L19 did not provide convincing explanation as to the deviation of data obtained from this large clone. It can be speculated that segments of sequence, which is highly similar to sequence from PAC p158L19, is present elsewhere in the genome, but this region has not yet been fully sequenced.

Although the vast majority of DGS/VCFS patients have typical 3 Mb deletions in 22q11, some display a nested distal deletion of 1.5 Mb. In both cases the breakpoints of the deletions are localized within the blocks of LCRs (Shaikh *et al.*, 2000). In addition, a few cases of atypical deletions have been reported in the literature both within and outside the TDR (O'Donnell *et al.*, 1997; Rauch *et al.*, 1999; Saitta *et al.*, 1999). There are also rare reports of DGS/VCFS patients, who do not display deletions in 22q11 at all; in some rare cases deletions have been observed in 10p13 (Daw *et al.*, 1996).

In two of our DGS/VCFS patients, we detected atypical, small deletions, encompassing one or two PAC clones (pac995o6 and pac699j1) (Figure 1, paper VI). Both latter PAC clones are located within the LCR-A block, which is the LCR containing the proximal breakpoint of typical deletions in the vast majority of DGS/VCFS patients. As mentioned above, in two cases no deletions were detected. These were two commercially obtained DGS/VCFS samples, where the diagnosis was made based on the clinical phenotype. No details could be obtained from the supplier regarding the clinical phenotype of these two patients. Today, the detection of deletions in DGS/VCFS patients is routinely performed by FISH analysis with two commercially available probes, N25 (D22S75) and TUPLE1 (D22S553) (Vysis, Inc). None of these probes would detect the small atypical deletions that were indicated by array analysis (Figure 1, paper VI). Although the true frequency of these atypical deletions in DGS/VCFS patients needs to be verified in a larger series of DGS/VCFS patients, these results are encouraging. We believe that array-CGH has the potential to increase the resolution of molecular diagnostics for this complex constellation of birth defects.

6 CONCLUDING REMARKS

The constitutional translocation t(11;22)(q23;q11) has caught the attention of several groups since it was first described. The fact that it is the most common reciprocal translocation occurring in constitutional DNA makes it interesting and worth studying. The cloning of the t(11;22)(q23;q11) breakpoints was not easy; the presence of low copy repeats (LCRs), high content of common repeats, inverted repeats and unclonable sequences made the characterization of the regions surrounding the breakpoints very challenging. Several reports, including our papers, have shown that the translocation breakpoints on both chromosomes occur in AT-rich regions, capable of forming palindromes, and that the breakpoints occur at the tip of these palindromes. In addition, the position of the breakpoints is highly conserved among carriers with different ethnic origins and with no common ancestor. Indeed, the 22q11 is a hotspot for chromosomal rearrangements, as it is also involved in the generation of deletions and duplications.

The exact mechanism of how the translocation arises is not fully understood. The fact that the breakpoints of another constitutional translocation t(17;22) have been shown to reside in the same location as for the t(11;22)(q23;q11) in chromosome 22 and in AT-rich palindromic sequences from chromosome 17 suggest that similar mechanisms underlie both translocations. All LCR-containing segments on 22q have been fully sequenced, except LCR-B containing the breakpoints of the t(11;22) translocation. The presence of these LCRs confer instability to the 22q11 region, this added to the fact that a large palindrome is formed in LCR-B is likely to be the cause of the unclonability of this region. With the sequences of the human genome in hand, it may be possible to identify additional chromosomal regions that are hotspots for rearrangements using similar mechanisms. Not much is known about the presence of palindromes in the human genome, but extensive studies in bacteria and yeast demonstrate that long DNA palindromes pose a threat to genome stability. It is possible that the formation of palindromic structures prevents the cloning of these sequences in vectors propagated in bacteria and may be the cause behind the more general unclonability of sequences in the human genome.

At the beginning, our interest in the characterization of the t(11;22) translocation breakpoints was based on the fact that a group of carriers have shown predisposition to develop breast cancer. One possible mechanism was that a gene (or genes) could be disrupted by the translocation and that this gene(s) could be responsible for the development of breast cancer in

some translocation carriers. Today we know that no genes seem to be disrupted by the translocation, in neither chromosome 11 nor 22. This, however, does not exclude a possibility that a positional effect on more distantly located genes may play role in the etiology of breast cancer in this group of translocation carriers.

Due to the high content of common repeats and other redundant sequences, the analysis of some regions in the human genome might be a very challenging task. One example is 22q11; the vast majority of patients with DGS/VCFS syndrome have deletions in this region. Many genes from 22q11 segment are deleted in DGS/VCFS patients and probably not all of them play role in the etiology of the syndrome. For instance, patients with similar deletions often display various degrees of disease severity. An improved molecular diagnosis may help in the further elucidation of which genes are involved in the development of this syndrome. By using different types of targets spotted on the array, we were able to construct the 22q11 array, specifically tailored to circumvent the problems of cross-hybridization caused by the common repeats and LCRs. At the same time, the resolution of analysis was greatly enhanced. These results encourage us to believe that most, if not all, medically relevant regions of our genome will eventually be accessible for analysis using high-resolution microarray-based CGH.

Several genomic disorders are caused by micro-deletions and micro-gains of chromosomal segments. With the current resolution offered by conventional methods for detection of gene dosage imbalances, such as FISH, it is evident that a majority of micro-deletions and micro-gains will escape detection. The development of the array-CGH methodology enables a high-resolution, comprehensive analysis of large genomic segments in the detection of gene dosage alterations. It is reasonably to assume that additional phenotypes will be linked to micro-deletions or micro-gains, when the high-resolution array-CGH (at the average level of 1 data point below 10 kb) is routinely applied to the analysis of the whole genome. The construction of our first array covering a human chromosome demonstrates that it will soon be possible to construct a whole genome microarray, with the average resolution at least as good as this shown here for 22q. This would allow a genome-wide screen, dramatically improving our possibilities for research and diagnostics of human diseases.

7 ACKNOWLEDGEMENTS

This thesis is a result of several years of work and many people has contributed to its results in a professional and personal way. I would like to acknowledge and express my gratitude

To Anna Wedell my supervisor for receiving me into her group in the middle time of this work, for all the scientific discussions, for her efficiency, kindness and patience.

To Jan Dumanski, my supervisor for all the scientific discussions and time that you spend on my education, for the thrust and for giving me the opportunity to develop as an independent student.

To Ingegerd Fransson, the person who was always there over these years, thank you very much for all the help in the lab and in practical matters, for all the coffee breaks, mushroom picking trips, for your friendship.

To all the colleagues at the Clinical Genetics Unit, and to Magnus Nordenskjöld, for the nice environment you provide to all of us. To the members in Anna's group, Tiina, Svetlana and Michela for the good atmosphere at the lab. Thanks to Tiina, my roommate at the lab for the nice time here and in Bertinoro.

To all the collaborators in my projects who made this work possible, at the CMM, MTC, Oklahoma University and Rudbeck Lab. Darek Kedra, Kevin O'Brien, Carl Bruder, Elisabeth Blennow, Maria Kost-Alimova, Stephan Imreh, Bruce Roe and Kiran Mantripragada.
To Sigrid Sahlén for all the innumerable FISH experiments you performed, and also to Ann-Christin Thelander, thanks for their excellent technical assistance. To Ann-Mari Dumanski and Yvonne Cowan for all the help with the administrative matters.

To the people in Jan Dumanskis former group, Kevin, Darek, Calle, Myriam, Eyal, Hoini, Ying Yang and Giedre for the nice times we had and also to all the members in the present Jan's group in Uppsala, for the enjoyable working atmosphere. Thanks to Kiran for all the time you spend with the DiGeorge array and in answering all my questions about the printing over and over again.

To Natalia, for your friendship and visits to Stockholm. I enjoy those very much.

To the ones who critically read this thesis, Anna Wedell, Elisabeth Blennow, Kevin O'Brien, Amilcar Flores, Gunther Weber and Jan Dumanski.

Thanks to all the people from Martin Schalling's former group at the floor 00, and to the girls from Gunther Weber's group Lovisa, Emma and Taranum for being so helpful when needed and for the nice atmosphere. To Lennart and Dagmar for the help with the computer problems.

To the Latin-American friends who passed and are at the CMM, Fabio, José, Juan, Amilcar, Roxana and Elizabeth. Thanks to Eliz for all the lunches, movies and chats we had.

To the people at the Department of Genetics, Stockholms University in Elisabeth Haggård and Ingrid Faye's group. To Virginia the person who introduced me into the lab work. To Ingrid, Yohannes, Kattis and Raúl, the group where I did my examination work for the good and friendly atmosphere.

To my Peruvian friends in Stockholm. To Cecilia my schoolmate I am very lucky to have you here. Thanks for all the trips we had together and for the support in the bad times as well as the good times. To Jose Carlos, Maria Victoria, Miguel, Hugo and Javier for being such a nice friends. To JC and Hugo for the additional "home support" in computer matters.
To Erik for the time we shared.

To my parents, for all love, care and worries. To my mother for always being around and her endless support. To my brothers Ernesto and Jorge Antonio and to my sister Maria Luisa, for being there. To my family in Peru for making me feel at home when I was there, despite all the years abroad and for keeping in touch.

To Micaela my daughter, the sweetest girl, for all the love and happiness you share.

8 REFERENCES

Bailey J. A., Yavor A. M., Viggiano L., Misceo D., Horvath J. E., Archidiacono N., Schwartz S., Rocchi M., and Eichler E. E. (2002). Human-specific duplication and mosaic transcripts: the recent paralogous structure of chromosome 22. *American Journal of Human Genetics* **70:** 83-100.

Beedgen B., Nutzenadel W., Querfeld U., and Weiss-Wichert P. (1986). "Partial trisomy 22 and 11" due to a paternal 11;22 translocation associated with Hirschsprung disease. *European Journal of Pediatrics* **145:** 229-232.

Beheshti B., Braude I.,Marrano, P., Thorner, P., Zielenska, M., Squire, J., (2003). Chromosomal localization of DNA amplifications in neuroblastoma tumors using cDNA microarray comparative genomic hybridization. *neoplasia* **In press**.

Blennow E., and Tillberg E. (1996). Small extra ring chromosome derived from chromosome 10p: clinical report and characterisation by FISH. *Journal of Medical Genetics* **33:** 399-402.

Bruder C. E., Ichimura K., Blennow E., Ikeuchi T., Yamaguchi T., Yuasa Y., Collins V. P., and Dumanski J. P. (1999). Severe phenotype of neurofibromatosis type 2 in a patient with a 7.4-MB constitutional deletion on chromosome 22: possible localization of a neurofibromatosis type 2 modifier gene? *Genes Chromosomes Cancer* **25:** 184-190.

Buckley P. G., Mantripragada K. K., Benetkiewicz M., Tapia-Paez I., Diaz De Stahl T., Rosenquist M., Ali H., Jarbo C., De Bustos C., Hirvela C., Sinder Wilen B., Fransson I., Thyr C., Johnsson B. I., Bruder C. E., Menzel U., Hergersberg M., Mandahl N., Blennow E., Wedell A., Beare D. M., Collins J. E., Dunham I., Albertson D., Pinkel D., Bastian B. C., Faruqi A. F., Lasken R. S., Ichimura K., Collins V. P., and Dumanski J. P. (2002). A full-coverage, high-resolution human chromosome 22 genomic microarray for clinical and research applications. *Human Molecular Genetics* **11:** 3221-3229.

Burn J. W., D.I. Cross, I. Atif, U. Scambler, P. Takao, A. Goodship, J. (1995). Developmental Mechanism of Heart Disease, Futura Publ. Armonk, New York.

Caspersson T., Farber S., Foley G. E., Kudynowski J., Modest E. J., Simonsson E., Wagh U., and Zech L. (1968). Chemical differentiation along metaphase chromosomes. *Experimental Cell Research* **49:** 219-222.

Caspersson T., Zech L., and Johansson C. (1970). Differential binding of alkylating fluorochromes in human chromosomes. *Experimental Cell Research* **60:** 315-319.

Chance P. F., Abbas N., Lensch M. W., Pentao L., Roa B. B., Patel P. I., and Lupski J. R. (1994). Two autosomal dominant neuropathies result from reciprocal DNA duplication/deletion of a region on chromosome 17. *Human Molecular Genetics* **3:** 223-228.

Chen K. S., Manian P., Koeuth T., Potocki L., Zhao Q., Chinault A. C., Lee C. C., and Lupski J. R. (1997). Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nature Genetics* **17:** 154-163.

Christian S. L., Fantes J. A., Mewborn S. K., Huang B., and Ledbetter D. H. (1999). Large genomic duplicons map to sites of instability in the Prader- Willi/Angelman syndrome chromosome region (15q11-q13). *Human Molecular Genetics* **8:** 1025-1037.

Collins J. E., Cole C. G., Smink L. J., Garrett C. L., Leversha M. A., Soderlund C. A., Maslen G. L., Everett L. A., Rice K. M., Coffey A. J., Gregory S., Gwilliam R., Dunham A., Davies A. F., Hassock S., Todd C. M., Lehrach H., Hulsebos T. J. M., Weissenbach J., Morrow

B., Kucherlapati R. S., Wadey R., Scambler P. J., Kim U.-J., Simon M. I., Peyrard M., Xie Y.-G., Carter N. P., Durbin R., Dumanski J. P., Bentley D. R., and Dunham I. (1995). A high resolution integrated yeast artificial chromosome clone map of human chromosome 22. *Nature* **377:** 367-379.

Daw S. C., Taylor C., Kraman M., Call K., Mao J., Schuffenhauer S., Meitinger T., Lipson T., Goodship J., and Scambler P. (1996). A common region of 10p deleted in DiGeorge and velocardiofacial syndromes. *Nature Genetics* **13:** 458-460.

Dean F. B., Hosono S., Fang L., Wu X., Faruqi A. F., Bray-Ward P., Sun Z., Zong Q., Du Y., Du J., Driscoll M., Song W., Kingsmore S. F., Egholm M., and Lasken R. S. (2002). Comprehensive human genome amplification using multiple displacement amplification. *Proceedings of the National Academy of Sciences U S A* **99:** 5261-5266.

Dean F. B., Nelson J. R., Giesler T. L., and Lasken R. S. (2001). Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Research* **11:** 1095-1099.

Dumanski J. P. (1996). The human chromosome 22-located genes and malignancies of the central nervous system. *Neuropathology and Applied Neurobiology* **22:** 412-417.

Dumanski J. P., Carlbom E., Collins V. P., Nordenskjold M., Emanuel B. S., Budarf M. L., McDermid H. E., Wolff R., O'Connell P., White R., and et al. (1991). A map of 22 loci on human chromosome 22. *Genomics* **11:** 709-719.

Dunham I., Shimizu N., Roe B. A., Chissoe S., Hunt A. R., Collins J. E., Bruskiewich R., Beare D. M., Clamp M., Smink L. J., Ainscough R., Almeida J. P., Babbage A., Bagguley C., Bailey J., Barlow K., Bates K. N., Beasley O., Bird C. P., Blakey S., Bridgeman A. M., Buck D., Burgess J., Burrill W. D., O'Brien K. P., and et al. (1999). The DNA sequence of human chromosome 22. *Nature* **402:** 489-495.

Edelmann L., Pandita R. K., and Morrow B. E. (1999a). Low-Copy Repeats Mediate the Common 3-Mb Deletion in Patients with Velo- cardio-facial Syndrome. *American Journal of Human Genetics* **64:** 1076-1086.

Edelmann L., Spiteri E., Koren K., Pulijaal V., Bialer M. G., Shanske A., Goldberg R., and Morrow B. E. (2000). AT-Rich Palindromes Mediate the Constitutional t(11;22) Translocation. *American Journal of Human Genetics* **68**: 1-13

Edelmann L., Spiteri E., McCain N., Goldberg R., Pandita R. K., Duong S., Fox J., Blumenthal D., Lalani S. R., Shaffer L. G., and Morrow B. E. (1999b). A common breakpoint on 11q23 in carriers of the constitutional t(11;22) translocation. *American Journal of Human Genetics* **65:** 1608-1616.

Evans D. G., Huson S. M., Donnai D., Neary W., Blair V., Newton V., Strachan T., and Harris R. (1992). A genetic study of type 2 neurofibromatosis in the United Kingdom. II. Guidelines for genetic counselling. *Journal of Medical Genetics* **29:** 847-852.

Feinberg A. P., and Vogelstein B. (1984). "A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity". Addendum. *Analytical Biochemistry* **137:** 266-267.

Fidlerova H., Senger G., Kost M., Sanseau P., and Sheer D. (1994). Two simple procedures for releasing chromatin from routinely fixed cells for fluorescence in situ hybridization. *Cytogenetics and Cell Genetics* **65:** 203-205.

Fraccaro M., Lindsten J., Ford C. E., and Iselius L. (1980). The 11q;22q translocation: a European collaborative analysis of 43 cases. *Human Genetics* **56:** 191-197.

Funke B., Edelmann L., McCain N., Pandita R. K., Ferreira J., Merscher S., Zohouri M., Cannizzaro L., Shanske A., and Morrow B. E. (1999). Der(22) syndrome and velo-cardio-

facial syndrome/DiGeorge syndrome share a 1.5-Mb region of overlap on chromosome 22q11. *American Journal of Human Genetics* **64:** 747-758.

Geschwind D. H., Gregg J., Boone K., Karrim J., Pawlikowska-Haddal A., Rao E., Ellison J., Ciccodicola A., D'Urso M., Woods R., Rappold G. A., Swerdloff R., and Nelson S. F. (1998). Klinefelter's syndrome as a model of anomalous cerebral laterality: testing gene dosage in the X chromosome pseudoautosomal region using a DNA microarray. *Developmental Genetics* **23:** 215-229.

Gordenin D. A., Lobachev K. S., Degtyareva N. P., Malkova A. L., Perkins E., and Resnick M. A. (1993). Inverted DNA repeats: a source of eukaryotic genomic instability. *Molecular Cell Biology* **13:** 5315-5322.

Heiskanen M. A., Bittner M. L., Chen Y., Khan J., Adler K. E., Trent J. M., and Meltzer P. S. (2000). Detection of gene amplification by genomic hybridization to cDNA microarrays. *Cancer Research* **60:** 799-802.

Hill A. S., Foot N. J., Chaplin T. L., and Young B. D. (2000). The most frequent constitutional translocation in humans, the t(11;22)(q23;q11) is due to a highly specific alu-mediated recombination. *Human Molecular Genetics* **9:** 1525-1532.

Hui A. B., Lo K. W., Yin X. L., Poon W. S., and Ng H. K. (2001). Detection of multiple gene amplifications in glioblastoma multiforme using array-based comparative genomic hybridization. *Laboratory Investigation* **81:** 717-723.

Iselius L., Lindsten J., Aurias A., Fraccaro M., Bastard C., Bottelli A. M., Bui T. H., Caufin D., Dalpra L., Delendi N., and et al. (1983). The 11q;22q translocation: a collaborative study of 20 new cases and analysis of 110 families. *Human Genetics* **64:** 343-55.

Jerome L. A., and Papaioannou V. E. (2001). DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nature Genetics* **27:** 286-291.

Kallioniemi A., Kallioniemi O. P., Sudar D., Rutovitz D., Gray J. W., Waldman F., and Pinkel D. (1992). Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* **258:** 818-821.

Kallioniemi O. P., Kallioniemi A., Sudar D., Rutovitz D., Gray J. W., Waldman F., and Pinkel D. (1993). Comparative genomic hybridization: a rapid new method for detecting and mapping DNA amplification in tumors. *Seminars in Cancer Biology* **4:** 41-46.

Kauraniemi P., Barlund M., Monni O., and Kallioniemi A. (2001). New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. *Cancer Research* **61:** 8235-8240.

Kurahashi H., and Emanuel B. S. (2001). Unexpectedly high rate of de novo constitutional t(11;22) translocations in sperm from normal males. *Nature Genetics* **29:** 139-40.

Kurahashi H., Shaikh T., Takata M., Toda T., and Emanuel B. S. (2003). The Constitutional t(17;22): Another Translocation Mediated by Palindromic AT-Rich Repeats. *American Journal of Human Genetics* **72:** 3.

Kurahashi H., Shaikh T. H., and Emanuel B. S. (2000a). Alu-mediated PCR artifacts and the constitutional t(11;22) breakpoint. *Human Molecular Genetics* **9:** 2727-2732.

Kurahashi H., Shaikh T. H., Hu P., Roe B. A., Emanuel B. S., and Budarf M. L. (2000b). Regions of genomic instability on 22q11 and 11q23 as the etiology for the recurrent constitutional t(11;22). *Human Molecular Genetics* **9:** 1665-70.

Laake K., Launonen V., Niederacher D., Gudlaugsdottir S., Seitz S., Rio P., Champeme M. H., Bieche I., Birnbaum D., White G., Sztan M., Sever N., Plummer S., Osorio A., Broeks A., Huusko P., Spurr N., Borg A., Cleton-Jansen A. M., van't Veer L., Benitez J., Casey G., Peterlin B., Olah E., Borresen-Dale A. L., and et al. (1999). Loss of heterozygosity at

11q23.1 and survival in breast cancer: results of a large European study. Breast Cancer Somatic Genetics Consortium. *Genes, Chromosomes & Cancer* **25:** 212-221.

Larsson C., Bystrom C., Skoog L., Rotstein S., and Nordenskjold M. (1990). Genomic alterations in human breast carcinomas. *Genes, Chromosomes & Cancer* **2:** 191-197.

Leach D. R. (1994). Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *Bioessays* **16:** 893-900.

Lindblom A., Sandelin K., Iselius L., Dumanski J., White I., Nordenskjold M., and Larsson C. (1994). Predisposition for breast cancer in carriers of constitutional translocation 11q;22q. *American Journal of Human Genetics* **54:** 871-876.

Lindsay E. A., Vitelli F., Su H., Morishima M., Huynh T., Pramparo T., Jurecic V., Ogunrinu G., Sutherland H. F., Scambler P. J., Bradley A., and Baldini A. (2001). Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* **410:** 97-101.

Lobachev K. S., Shor B. M., Tran H. T., Taylor W., Keen J. D., Resnick M. A., and Gordenin D. A. (1998). Factors affecting inverted repeat stimulation of recombination and deletion in Saccharomyces cerevisiae. *Genetics* **148:** 1507-1524.

Merscher S., Funke B., Epstein J. A., Heyer J., Puech A., Lu M. M., Xavier R. J., Demay M. B., Russell R. G., Factor S., Tokooya K., Jore B. S., Lopez M., Pandita R. K., Lia M., Carrion D., Xu H., Schorle H., Kobler J. B., Scambler P., Wynshaw-Boris A., Skoultchi A. I., Morrow B. E., and Kucherlapati R. (2001). TBX1 is responsible for cardiovascular defects in velo-cardio- facial/DiGeorge syndrome. *Cell* **104:** 619-629.

O'Donnell H., McKeown C., Gould C., Morrow B., and Scambler P. (1997). Detection of an atypical 22q11 deletion that has no overlap with the DiGeorge syndrome critical region. *American Journal of Human Genetics* **60:** 1544-1548.

Pinkel D., Segraves R., Sudar D., Clark S., Poole I., Kowbel D., Collins C., Kuo W. L., Chen C., Zhai Y., Dairkee S. H., Ljung B. M., Gray J. W., and Albertson D. G. (1998). High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nature Genetics* **20:** 207-211.

Pollack J. R., Perou C. M., Alizadeh A. A., Eisen M. B., Pergamenschikov A., Williams C. F., Jeffrey S. S., Botstein D., and Brown P. O. (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nature Genetics* **23:** 41-46.

Rauch A., Pfeiffer R. A., Leipold G., Singer H., Tigges M., and Hofbeck M. (1999). A novel 22q11.2 microdeletion in DiGeorge syndrome. *American Journal of Human Genetics* **64:** 659-666.

Rouleau G. A., Merel P., Lutchman M., Sanson M., Zucman J., Marineau C., Hoang-Xuan K., Demczuk S., Desmaze C., Plougastel B., and et al. (1993). Alteration in a new gene encoding a putative membrane-organizing protein causes neuro-fibromatosis type 2. *Nature* **363:** 515-521.

Ruttledge M. H., Andermann A. A., Phelan C. M., Claudio J. O., Han F. Y., Chretien N., Rangaratnam S., MacCollin M., Short P., Parry D., Michels V., Riccardi V. M., Weksberg R., Kitamura K., Bradburn J. M., Hall B. D., Propping P., and Rouleau G. A. (1996). Type of mutation in the neurofibromatosis type 2 gene (NF2) frequently determines severity of disease. *American Journal of Human Genetics* **59:** 331-342.

Saitta S. C., McGrath J. M., Mensch H., Shaikh T. H., Zackai E. H., and Emanuel B. S. (1999). A 22q11.2 deletion that excludes UFD1L and CDC45L in a patient with conotruncal and craniofacial defects. *American Journal of Human Genetics* **65:** 562-566.

Sambrook J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning; a laboratory manual. *Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press.*

Sanger F., Nicklen S., and Coulson A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences U S A* **74:** 5463-5467.

Scambler P. J. (2000). The 22q11 deletion syndromes. *Hum Mol Genet* **9:** 2421-2426.

Shaikh T. H., Budarf M. L., Celle L., Zackai E. H., and Emanuel B. S. (1999). Clustered 11q23 and 22q11 breakpoints and 3:1 meiotic malsegregation in multiple unrelated t(11;22) families. *American Journal of Human Genetics* **65:** 1595-1607.

Shaikh T. H., Kurahashi H., Saitta S. C., O'Hare A. M., Hu P., Roe B. A., Driscoll D. A., McDonald-McGinn D. M., Zackai E. H., Budarf M. L., and Emanuel B. S. (2000). Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Human Molecular Genetics* **9:** 489-501.

Solinas-Toldo S., Lampel S., Stilgenbauer S., Nickolenko J., Benner A., Dohner H., Cremer T., and Lichter P. (1997). Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* **20:** 399-407.

Sonnhammer E. L., and Durbin R. (1995). A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* **167:** GC1-10.

Staden R., Ed. (1994). "The Staden package," Humana Press Inc, Totawa, NJ.

Stalmans I., Lambrechts D., De Smet F., Jansen S., Wang J., Maity S., Kneer P., Von Der Ohe M., Swillen A., Maes C., Gewillig M., Molin D. G., Hellings P., Boetel T., Haardt M., Compernolle V., Dewerchin M., Plaisance S., Vlietinck R., Emanuel B., Gittenberger-De Groot A. C., Scambler P., Morrow B., Driscol D. A., Moons L., Esguerra C. V., Carmeliet G., Behn-Krappa A., Devriendt K., Collen D., Conway S. J., and Carmeliet P. (2003). VEGF: A modifier of the del22q11 (DiGeorge) syndrome? *Nature Medicine* **9:** 173-182.

Stankiewicz P., and Lupski J. R. (2002). Molecular-evolutionary mechanisms for genomic disorders. *Current Opinion in Genetics and Development* **12:** 312-319.

Tapia-Paez I., Kost-Alimova M., Hu P., Roe B. A., Blennow E., Fedorova L., Imreh S., and Dumanski J. P. (2001). The position of t(11;22)(q23;q11) constitutional translocation breakpoint is conserved among its carriers. *Human Genetics* **109:** 167-177.

Trask B. J. (1991). Fluorescence in situ hybridization: applications in cytogenetics and gene mapping. *Trends in Genetics* **7:** 149-154.

van Ommen G. J., Breuning M. H., and Raap A. K. (1995). FISH in genome research and molecular diagnostics. *Current Opinion in Genetics and Development* **5:** 304-308.

Xie Y. G., Han F. Y., Bajalica S., Blennow E., Kristoffersson U., Dumanski J. P., and Nordenskjold M. (1994). Identification, characterisation and clinical applications of cosmids from the telomeric and centromeric regions of the long arm of chromosome 22. *Human Genetics* **94:** 339-345.

Zackai E. H., and Emanuel B. S. (1980). Site-specific reciprocal translocation, t(11;22) (q23;q11), in several unrelated families with 3:1 meiotic disjunction. *American Journal of Human Genetics* **7:** 507-521.

Zhou Z. H., Akgun E., and Jasin M. (2001). Repeat expansion by homologous recombination in the mouse germ line at palindromic sequences. *Proceedings of the National Academy of Sciences U S A* **98:** 8326-8333.

Zucman-Rossi J., Legoix P., Der Sarkissian H., Cheret G., Sor F., Bernardi A., Cazes L., Giraud S., Ollagnon E., Lenoir G., and Thomas G. (1998). NF2 gene in neurofibromatosis type 2 patients. *Human Molecular Genetics* **7:** 2095-2101.