1 Introduction

1.1 Structural abnormalities

Chromosomal abnormalities and gene mutations are some of the major causes of human sufferings. Structural changes in chromosomes can be classified according to cytological types and their effect on the phenotype. The main cytological types are translocations, deletions, inversions, insertions, isochromosomes, dicentric chromosomes and ring chromosomes (Fig. 1). Structural rearrangements alter the genomic architecture and may result in human disease traits. These alterations can be termed as Genomic disorders. Genomic disorders can lead to the complete loss or gain of a gene(s) sensitive to a dosage effect, or can disrupt the structural integrity of a gene (Lupski, 1998). Genome alterations can occur through many mechanisms, one of which is the nonallelic homologous recombination during meiosis between region specific, low-copy repeated sequences (Weinstock et al., 1998). Recombination between direct repeats leads to deletions and duplications, while between inverted repeats results in inversion of the intervening genomic sequences.

1.1.1 Chromosomal rearrangements as tools for disease gene identification

Balanced chromosomal rearrangements, which truncate, delete or otherwise inactivate genes, are powerful tools for mapping novel disease genes (Tommerup et al., 1993; Collins, 1995; Bugge et al., 2000). The cytogenetic and molecular analyses of such rearrangements constitute an efficient strategy for mapping and cloning these disease genes. There are many examples of human diseases in which analyses of chromosomal anomalies have led to the discovery of diseases and further identification of genes involved in the etiology of the disease. A few among those are X-linked ectodermal dysplasia (Srivastava et al., 1996), Aniridia (Gessler et al., 1989) and holoprosencephaly (Belloni et al., 1996).
Introduction

Structural chromosomal abnormalities involve chromosome breakage.

- **One break on one chromosome**
  - Terminal deletion
    - Acentric fragment is lost

- **Two breaks on one chromosome**
  - Reciprocal Translocation
    - Balanced exchange of acentric fragments

- **Two breaks on different chromosomes**
  - Interstitial deletion
    - Region between breakpoints is discarded; terminal fragments fuse
  - Centric translocation
    - Fusion of centric fragments of two acrocentric chromosomes

- **Three breaks at least two on one chromosome**
  - Insertional translocation
    - Region between two breakpoints on one chromosome is excised and inserts at location of third breakpoint which may be on the same or a different chromosome

Fig. 1: Major types of structural chromosomal abnormalities (adapted from Strachan and Read, 1996).

In majority of the cases a single gene is expected to be disrupted either in its structure or expression, thus leading to a disease phenotype. An autosomal dominant disorder can be caused either by a gain-of-function mutation on one of the alleles or by the loss-of-function on one homologue (haploinsufficiency). Chromosomal rearrangements associated with an autosomal dominant disorder are usually expected to cause haploinsufficiency, since the disruption of the coding region of a gene may lead to a functional-null-allele (Vortkamp et al., 1991; Fahnold et al., 1995). In most of the cases chromosomal aberrations occur sporadically (de novo) and do not segregate within a family. Such cases are simple in interpreting the correlation of the patient phenotype with the chromosomal aberration. In rare cases chromosomal aberrations segregate within a family with a specific clinical phenotype. Such rare families with chromosomal rearrangements segregating with an “autosomal dominant” trait are highly significant for identifying the disease causing gene(s). In most such cases one of the chromosomal
breakpoints map within or close to the respective gene thus disturbing its gene activity. If a gene is located in the breakpoint region that is disrupted then it can be considered as a candidate gene for the disease observed in the patient. Mutation screening of this candidate gene can be done in patients presenting with the same disorder and having a normal karyotype. The identification of mutations in this gene will thus establish a causal relation between the mutation and that particular disease. For example Tsukamoto et al. (1992) cloned and characterized an inversion breakpoint on chromosome 2q35 with Waardenburg syndrome type I. The HuP2 gene is found to be disrupted by the inversion and thus suggested that this gene is a candidate for the Waardenburg syndrome type I. Hence, the identification of disease causing genes is a prerequisite for proper diagnosis and genetic counselling of patients and their families as well as for studying the molecular pathogenesis leading to the disease.

1.1.2 Mechanism of inversion

Two breaks in a single chromosome can lead to an inversion, or an interstitial deletion or else a ring chromosome. An inversion results when the chromosomal segment between the two breakpoints takes a 180 degree turn prior to resealing the breakpoints (Fig. 2). Usually, this is a balanced rearrangement with no loss of chromosomal material. The inverted segment may include the centromere (pericentric inversion) or may be confined to one chromosome arm (paracentric inversion).

![Fig. 2: Schematic representation showing the mechanism of inversion.](image)
1.2 The Human Genome Project

The Human Genome Project (HGP) is an international project conceived in 1986, which officially began in 1990 and was completed in 2003. The complete nucleotide sequence of the human genome was one of the several goals of the HGP. The major goal then was to acquire the fundamental information concerning our genetic make-up. This is crucial for basic scientific understanding of human genetics and the role of various genes in health and disease. The detail interaction with research focussing on mapping disease genes is shown in Figure 3. Knowledge of the human genome sequence enables us to understand how the genetic information determines the development, the structure and the function of the human body. It helps in exploring how variations in our human genome can cause disease, how they affect our interaction with our environment, and more importantly how to develop new and effective ways to improve human health (Bentley, 2000).

![Diagram of major scientific strategies and approaches used in the Human Genome Project](image)

**Key:**
- Inputs
- Physical methods
- Genetic methods
- Black cloud

Fig. 3: Major scientific strategies and approaches used in the Human Genome Project (adapted from Strachan and Read, 1996)
Physical map of the Human genome

A physical map is not only a scaffold for genomic sequencing but also offers access to any genomic region, which is very important for gene cloning. Gene maps have been constructed in humans and different mammalian species for two reasons: first as a resource for locating the genetic determinants of heritable characters, behaviour and phenotypes; and second as a template for resolving and interpreting pattern of evolving genome organization in their ancestry. A new dimension has been introduced into biomedical research by HGP and other related programmes for various organisms. The main goal is to determine the sequence of three billion nucleotide pairs in the DNA of the human genome and to find all the genes therein.

A variety of different physical maps of the human genome like Cytogenetic, Chromosome breakpoint maps, Restriction map, Clone contig map, Sequence-tagged site (STS) map, Expressed sequence tag (EST) map and DNA sequence maps, were constructed and further are used to map the human genome.

The first physical map was obtained when cytogenetic banding technique made way for subchromosomal identification (ISCN, 1995). Though the resolution was low, it has been very useful as a framework for assigning the locations of human DNA sequences by chromosome in situ hybridization techniques. Some rare-cutter restriction maps have been achieved for a few human chromosomes; and the first restriction map using NotI restriction enzyme on chromosome 21 was published by Ichikawa et al. (1993). With the advent of the YAC system (Burke et al., 1987) it became possible to clone large inserts up to 2 Mb (Anderson, 1993) and the first generation physical map of the human genome was constructed with CEPH YAC library by Cohen et al. (1993). But a major disadvantage of the YAC libraries is its high frequency of chimeric clone formation (Green et al., 1991; Bray-Ward et al., 1996). This problem was overcome to some extent by verifying the physical location of the individual YAC by FISH and some were established in terms of cytogenetic bands on some chromosomes. Since most of the YAC clones were positive for STS markers (Hudson et al., 1995) with defined genetic linkage distances (Dib et al., 1996; Broman et al., 1998) corresponding to their physical locations, the data facilitated integration of cytogenetic, genetic and physical maps of the human genome (Bray-Ward et al., 1996; Cox et al., 1996). This integration has been particularly useful in the maximum coverage of the genome (Crollius et al., 1996; Bouffard et al.,
However in some gene rich regions, higher resolution contig maps were constructed using overlapping cosmid, bacteriophage lambda, P, P1-artificial chromosome (PAC) and Bacterial artificial chromosome (BAC) clones. The development of cloning systems like PAC (Ioannou et al., 1994) and BAC (Shizuya et al., 1992) proved to be quite important in the success of the whole-genome map as they were more stable than YAC clones. Ultimately, high resolution maps based on PAC and BAC clones provided suitable framework for sequencing and much better coverage of the genome. Though initial efforts to construct clone based regional and even chromosomal physical maps of the human genome using cosmid libraries derived from isolated human chromosomes met with limited success (Doggett et al., 1995), a reasonable physical map of the human genome based on clones was published in 2001 by McPherson et al. The physical map of the overlapping clones was constructed by checking all BACs from different libraries for the restriction fragments pattern, and thus established a fingerprint for each BAC (Soderlund et al., 1997; Sulston et al., 1989). This helped in distinguishing different BACs, and in turn assessing the degree of overlaps. Hence, the clone based map has been vital for the accurate assembly of the human genome sequence without any gaps (International Human Genome Mapping Consortium (IHGMC), 2001). Other types of map like the EST map required cDNA sequencing and also mapping cDNAs back to other physical maps (Adams et al., 1995). Subsequently, subchromosomal localization of human cDNA clones by FISH has also been possible (Korenberg et al., 1995). A complete contig map of a chromosome would therefore comprise the entire DNA without any gaps. Finally, obtaining the physical map with complete nucleotide sequence of the genome providing the highest possible resolution up to a single base pair is the desired ultimate goal.

The advances of the HGP and the completion of the total genome sequences of many species enable us to view the gene information of the entire genome. As a result the mechanisms for some of the genetic diseases are best understood at a genomic level.
1.2.1 Identification of human disease genes

The resources provided by the HGP are shaping the strategies used for disease gene identification (Ballabio, 1993). The identification of genes involved in human diseases is important to understand the pathophysiology of the disease, further it often provides new insights into normal human development and biology. Basically, diseases can be identified by four main approaches. They are positional cloning, functional cloning, positional candidate approach and cloning of a candidate gene. A general strategy of functional cloning requires prior knowledge of the function of the gene. For example the gene for phenylketonuria was identified by purifying the enzyme from liver and raising antibodies which immunoprecipitated polysomes containing phenylalanine-hydroxylase mRNA (Robson et al., 1982). Positional cloning, a term coined by Collins in 1992, requires the knowledge of its physical location in the genome and usually with little information about the function. This requires laborious methods of chromosome walking and the identification of expressed sequences. The first gene isolated by positional cloning was chronic granulomatous disease gene in 1986 by Royer-Pokora et al. The genes for many important diseases like Duchenne muscular dystrophy, Cystic fibrosis, Huntington’s disease, Adult polycystic kidney disease, Colorectal cancer, and Breast cancer were also isolated by positional cloning (Ballabio, 1993). The candidate gene approach relies on the availability of the information from previously isolated genes (Collins in 1992). The positional candidate approach involves the combination of mapping the genes to the correct chromosomal subregion and then checking for candidates in that region. For example Marfan syndrome was mapped to chromosome 15q using the information available from both positional and candidate approach (Kainulainen et al., 1990; Magenis et al., 1991; Dietz et al., 1991).
1.3 Subject of the present study

In this work, a three generation familial pericentric inversion of chromosome 3 with short stature is studied. The proband had been referred to the Department of Human Genetics and Medical Biology, Martin-Luther University, Halle-Wittenberg for chromosomal analysis of short stature. She was born in 1986 and was 146 cm tall in 2002. Apart from short stature, no other dysmorphic features were identified in either clinical or biochemical point of view. Initial cytogenetic analysis revealed a karyotype of 46, XX inv (3)(p23; q25q26)(Fig. 4A). To establish the origin of this pericentric inversion, chromosomal analyses of the parents and the sibling were done. The karyotypes of both the father and the brother were normal with 46,XY but the karyotype of the mother showed the same inversion as that of the proband. In order to further elucidate the inversion, the chromosomes of the maternal grandmother and her other daughter were also analyzed. The maternal aunt of the proband showed a normal karyotype, whereas the maternal grandmother carried the same inversion. The pedigree of the family, along with the year of birth, heights and the initial karyotypes in three generations, is shown in Figure 4B.

Fig. 4A: The GTG banded partial karyogram of the proband showing the pericentric inversion of chromosome 3 and the ideogram showing the cytogenetically mapped inversion breakpoint regions.
1.4 Possible causes of short stature

The causes of short stature are heterogeneous. Many genes and environmental influences are involved in the individual shaping of body size. So a systematic classification of the various syndromes involving short stature is not exactly interpreted till date (Enders, 1992).

Hormonal involvement

Growth hormone deficiency may occur by itself or in combination with one or more other pituitary hormone deficiencies. Growth hormone (GH) deficiency is a rare cause of short stature. The gene encoding growth hormone is GH1, which shows a variety of defects, including deletions, frameshifts, splice sites and nonsense mutations, and leads to either decreased GH expression or action (Wagner et al., 1998). A novel homozygous 5’ splice site mutation in the Growth Hormone-Releasing hormone Receptor (GHRHR) gene as cause of dwarfism with familial isolated growth hormone deficiency (IGHD) was
described by Salvatori et al. (1999). In pituitary dwarfism, the Pit-1 gene encodes the POU-domain transcription factor Pit-1 which is important for the development of the anterior pituitary gland and the expression of the GH, prolactin (PRL) and Thyroid stimulating hormone (TSH) genes. R271W mutation of the Pit-1 gene causes hypoplasia of the pituitary gland and deficiencies of those hormones (Aarskog et al., 1997). Many mutations have been identified in the transcription factors which regulate the pituitary development like Pit1 (POU1F1) (Pfaeffle et al., 1992), PROP1 (Wu et al., 1998), HESX1 (Dattani et al., 1998) or LHX3 (Netchine et al., 2000) and LHX4 (Machinis et al., 2001).

**Major syndromes involved in short stature**

About 74 major syndromes describe short stature as one of the characteristic features (OMIM). However, only a few syndromes are described here. SHORT syndrome associated with mild intrauterine growth retardation, postnatal growth deficiency and delayed bone age was first reported in 1975 by Gorlin et al. and Sensenbrenner et al. Achondroplasia (ACH) is the common genetic form of dwarfism. ACH candidate region includes the gene encoding fibroblast growth factor receptor 3 (FGFR3). Point mutation in the FGFR3 gene in ACH heterozygotes and homozygotes is responsible for ACH (Shiang et al., 1994). The most common cytogenetic cause of short stature in females is Turner syndrome (Ullrich, 1930; Turner, 1938). It is a common developmental disorder in females which is also characterized by short stature. A pseudoautosomal location for a dosage-sensitive locus involved in stature has been suggested based on the analyses with deletions of a specific segment of the short arm pseudoautosomal region (PAR1); hemizygosity for this putative locus probably contributes to the short stature in Turner individuals. Individuals with deletions of portions of short arm of PAR1 were short (Ogata et al., 1992). Two allelic variants of Turner syndrome Leri-Weill Dyschondrosteosis and Langer mesomelic dysplasia were also reported (Belin et al., 1998; Shears et al., 1998). Dyschondrosteosis is an autosomal dominant form of mesomelic dysplasia where short stature due to shortening of the forelegs, madelung deformity of the forearm with bowing of the radius and dorsal dislocation of the distal ulna is frequently observed.
Identification of genes influencing short stature

**On sex chromosomes**

The genetic causes of growth abnormalities were well reviewed by Mendez et al. in 1985. The identification of genes that influence the complex trait height is more complex. However, candidate gene and positional cloning approaches, using linkage analysis on a whole genome scale, helped in isolating some genes on both sex and autosomal chromosomes. Many chromosomal aberrations in patients with short stature were reported. The Short stature HoMeobox-containing gene (SHOX) on the short arm of the X and Y chromosome is an important determining factor of stature phenotype (Musebeck et al., 2000). The SHOX gene was mapped to the Xp22.3 (Ellison et al., 1997) and the haploinsufficiency of this gene leads to growth failure and short stature (Rao et al., 1997). Interestingly, SHOX mutations have also been described as causative for the Leri-Weill syndrome, a mesomelic short stature syndrome (Belin et al., 1998). In addition, homozygous SHOX mutations have been shown to cause Langer-type mesomelic dwarfism (Belin et al., 1998; Shears et al., 1998). A homeodomain identical to SHOX called PHOG was found to be a candidate for involvement in the short stature of Turner syndrome (Ellison et al., 1997).

**On autosomal chromosomes**

Autosomal aberrations with short stature also have been reported. The SHOX Homologous gene on chromosome Three (SHOT) a SHOX-related human gene has been identified and localized on chromosome 3q25-26 (Blaschke et al; Baere et al., 1998). This SHOT is closely related to SHOX on sex chromosomes and also has a murine counterpart OG12X (Clement-Jones et al., 2000). Some of the known growth related genes identified on other autosomes are Cathepsin K on 1p13, FGFR3 on 4p16, Oestrogen receptor on 6q25, GHRH receptor on 7p14, ROR2 on 9q22, LHX3 on 9q32-34, Col1A1, GH1 on 17q22-23 (Pfaeffle, 2006), and Growth hormone-releasing factor (GHRF) or Somatocrinin on 20p12 (Rao et al., 1991).

1.5 **Insight into some common chromosomal rearrangements involving 3q26**

Many different structural anomalies involving 3q26 region are well reported in the literature. The structural abnormalities taking the form of intra- and inter-chromosomal
rearrangements either between chromosomes 3, for example t(3;3)(q21;q26) or between
3q21 or 3q26 and other chromosomes were reported by Pintado et al., 1985. But majority
of the rearrangements with 3q26 are involved with cancer. For example translocations or
inversions of chromosome 3 with breakpoints involving band 3q26 were specifically
associated with megakaryoblastic acute phase or abnormal megakaryocytopoiesis. The
occurrence in myelodysplastic syndromes and myeloid leukemias of acquired
abnormalities of chromosome 3 is well documented by Belloma et al. (1992). Acute
leukaemia with abnormal thrombopoiesis is also associated with pericentric inversion or
homologous translocations of chromosome 3 involving bands 3q21 and 3q26.2; these
were described in a few patients with acute non lymphocytic leukemias (ANLL) (Bitter et
al., 1985) and dysmyelopoietic syndromes (DMPS) (Carroll et al., 1986) or accelerated
phases of chronic granulocytic leukaemia (CGL) (Carbonell et al., 1982). This
association suggested the presence of important thrombopoiesis gene on chromosome 3q
(Bitter et al., 1985). Ecotropic virus integration-1 (EVI1) gene located at 3q26 has been
reported in individuals with AML involving translocation or inversions of chromosome 3
long arm (Levy et al., 1994). Philadelphia chromosome as a secondary abnormality in
inv(3)(q21q26) acute myeloid leukaemia (Han and Theil, 2006) and t(3;17)(q26;q22) as
an additional change in Philadelphia positive chronic myelogenous Leukemia in
acceleration were also observed (Mugneret et al., 1992).

1.6 Chromosome 3 and short stature association
A few cases associated with short stature and chromosome 3 abnormalities are also
known in the literature. For example, a case of Turner syndrome with a familial
pericentric inversion inv (3)(p25q21)pat. This case showed primarily short stature with
minor form of Turner syndrome like monosomy and isochromosome X and a familial
pericentric inversion on chromosome 3 (Stine et al., 1982). Also, a female child with
short stature showing ring chromosome 3 abnormalities along with hypoplastic thumb
and coloboma of iris (Barajas et al., 2001) was described. Two cases of trisomy 21(pter-
q22.1) with no major features of Down syndrome but with moderate mental retardation
and short stature were also reported. Interestingly, the extra chromosome 21 contained
the distal part of chromosome 3p in one case and 14q in the other (Kondo et al., 2006). In
another case a 11 year old boy with short stature and learning difficulties showed a
chromosome 3p23 break with ring formation and translocation of displaced 3p23-->pter segment to 6pter (Yip et al., 1996). A case of distal segment of chromosome 3q deletion with primary ovarian failure and short stature as one of the major feature was reported by Nguyen et al. (2005). Finally, a partial monosomy 3q in a boy with short stature, developmental delay and mild dysmorphic features was also reported (Brueton et al., 1989).

1.7 Aim of the study

As a primary aim, an inversion chromosome 3 segregating in a three generation family with short stature was studied by positional cloning:

- Initially, YAC (Yeast Artificial Chromosome) clones should be selected by \textit{in silico} analysis using the human genome database and should be used to delineate the inversion breakpoint by FISH (Fluorescence \textit{in situ} hybridization).
- Further, BAC (Bacterial Artificial Chromosome) clones should be selected and breakpoints spanning clones on both 3p and 3q regions should be identified and characterization of those clones by different strategies and techniques should be used to narrow down the breakpoint.
- Data mining using the DNA sequences of the clones anchoring the breakpoints should be carried out to find out known or putative gene(s) located in or close to the critical regions which might be a candidate gene for this phenotype.
- By molecular methods the junction fragments of the breakpoint regions should be characterized.
- Subsequent \textit{in silico} analysis using the public databases should give information about the gene contents and the sequence features, like repetitive elements and segmental duplications at the inversion breakpoint region. The characterization of the DNA motifs at the breakpoint regions should also be performed to understand the mechanism underlying this familial inversion.

The detailed analysis of the inversion breakpoints might help to provide a better understanding of the reasons causing short stature, such as the forces driving chromosomal rearrangements and structural types promoting breakage, but also may provide insights in human pathology of short stature.