A Physical Map of the Short Arm/Centromere Region of Human Chromosome 21

Jennifer A. Doll
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_theses

Part of the Biology Commons

Recommended Citation
https://ecommons.luc.edu/luc_theses/4040

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1994 Jennifer A. Doll
LOYOLA UNIVERSITY OF CHICAGO

A PHYSICAL MAP OF THE SHORT ARM/CENTROMERE REGION OF HUMAN CHROMOSOME 21

A THESIS SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL IN CANDIDACY FOR THE DEGREE OF MASTER OF SCIENCE DEPARTMENT OF BIOLOGY

BY

JENNIFER A. DOLL

CHICAGO, ILLINOIS

JANUARY 1994
Chromosome 21 serves as a useful model to study nondisjunction and translocation in human cells because of its frequent and well-documented involvement in the nondisjunction- and translocation-associated conditions such as Down Syndrome, Alzheimer Disease, premature centromere division, and cancers. The repetitive sequences present on the short arm/centromere of the chromosome have specifically been implicated in such nondisjunction and translocation events. The specific purpose of this project was to construct a physical map for the short arm/centromeric region of chromosome 21 using pulsed field gel electrophoresis and hamster-human hybrid cell lines containing chromosome 21 or fragments of it as the sole human component. Five probes representing different tandem repetitive sequence families were utilized in the study: pHH550-31 and pHE340-9 (alphoid family probes), Kpn1.8 (a satellite III sequence), H7-1 (a β-satellite sequence), and CHB (a ribosomal gene probe).

Non-overlapping restriction fragments containing the repetitive sequences studied, including the ribosomal genes,
totaled 21.2 to 31.9 Mb, well over the approximately 20 Mb originally estimated for the size of the short arm. Linkage was established between the two alphoid sequences, with approximately two Mb separating the two families. Multiple linkage groups were established between the satellite III and β-satellite sequences, both proximal and distal to the rDNA. Satellite III sequences were observed to span the breakpoint of the 153E7BX cell line. A cluster of HH550-31 alphoid sequences and a satellite III cluster are both present in the centromeric region of the chromosome. This type of organization is similar to that found at other human centromeres. A multiple cluster organization for all of the families studied, with the exception of the HE340-9 alphoid family, was evident from the data obtained. This type of organization is consistent with out-of-register recombination models that could explain nondisjunction or translocation events.
ACKNOWLEDGEMENTS

I would like to thank all my Chicago friends for their support during this step in my life. I would especially like to thank my advisor, Dr. Jeffrey Doering, for all his support and guidance. I also thank my committee members, Drs. John Smarrelli and Holden Maecker of Loyola University and Dr. Michael Cummings of The University of Illinois at Chicago, for their suggestions and comments. I would also like to thank everyone in the Doering lab for their help and humor, especially Sheryl Cammarata, MaryAnn Lee, Marissa Michaels and Annie Burket. I also am indebted to Susan Guggenheim for her help in raising my computer literacy level. And, of course, I could not forget to thank Joe Schluep and Dr. Warren Jones for all their help (and sarcasm) in the dark room.

I am especially indebted to all my friends here in Chicago who helped me keep my sanity during this time. I could not have done it without you. I would also like to thank my parents for everything they have done for me. I greatly appreciate all the sacrifices you have made for me, and I dedicate my thesis to you both.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>REVIEW OF RELATED LITERATURE</td>
<td>5</td>
</tr>
<tr>
<td>Down Syndrome</td>
<td>6</td>
</tr>
<tr>
<td>Alzheimer Disease</td>
<td>9</td>
</tr>
<tr>
<td>Premature Centromere Division</td>
<td>11</td>
</tr>
<tr>
<td>Cancer</td>
<td>12</td>
</tr>
<tr>
<td>Acrocentric Chromosomes</td>
<td>15</td>
</tr>
<tr>
<td>Chromosome 21</td>
<td>17</td>
</tr>
<tr>
<td>Repetitive DNA</td>
<td>19</td>
</tr>
<tr>
<td>Repetitive Sequences Localized on Chromosome 21</td>
<td>22</td>
</tr>
<tr>
<td>DNA Sequence Probes Utilized in this Study</td>
<td>24</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>29</td>
</tr>
<tr>
<td>Experimental Approach</td>
<td>29</td>
</tr>
<tr>
<td>DNA Sources</td>
<td>31</td>
</tr>
<tr>
<td>Preparation of DNA Samples</td>
<td>31</td>
</tr>
<tr>
<td>Restriction Enzyme Digests</td>
<td>33</td>
</tr>
<tr>
<td>TAFE Gel Conditions</td>
<td>34</td>
</tr>
<tr>
<td>Transfer and Hybridization</td>
<td>36</td>
</tr>
<tr>
<td>RESULTS</td>
<td>39</td>
</tr>
<tr>
<td>Determination of Restriction Fragments</td>
<td>39</td>
</tr>
<tr>
<td>Sequence Localization</td>
<td>43</td>
</tr>
<tr>
<td>Linkage</td>
<td>45</td>
</tr>
<tr>
<td>Double Digests of 153E7BX probed with HH550-31</td>
<td>48</td>
</tr>
<tr>
<td>Ribosomal Probe</td>
<td>49</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>70</td>
</tr>
<tr>
<td>Construction of the Physical Map</td>
<td>71</td>
</tr>
<tr>
<td>Description of the Physical Map</td>
<td>77</td>
</tr>
<tr>
<td>Implications of the Organization Pattern</td>
<td>81</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.</td>
<td>Autoradiograms of restriction digests hybridized with the HH550-31 alphoid probe</td>
</tr>
<tr>
<td>2.</td>
<td>Autoradiograms of restriction digests hybridized with the HE340-9 alphoid probe</td>
</tr>
<tr>
<td>3.</td>
<td>Autoradiograms of restriction digests hybridized with the β-satellite probe</td>
</tr>
<tr>
<td>4.</td>
<td>Autoradiograms of restriction digests hybridized with the satellite III probe</td>
</tr>
<tr>
<td>5.</td>
<td>Diagram of chromosome 21 fragments in the somatic cell hybrids used in this study</td>
</tr>
<tr>
<td>6.</td>
<td>Demonstration of linkage between the two alphoid family sequences</td>
</tr>
<tr>
<td>7.</td>
<td>Demonstration of linkage between the satellite III and β-satellite sequences</td>
</tr>
<tr>
<td>8.</td>
<td>Demonstration of non-linkage between the alphoid and non-alphoid sequences</td>
</tr>
<tr>
<td>9.</td>
<td>Autoradiograms of restriction digests hybridized with the ribosomal DNA probe</td>
</tr>
<tr>
<td>10.</td>
<td>A partial physical map of the short arm/centromere region of chromosome 21</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.</td>
<td>Characteristics of the PFGE programs used in this study</td>
</tr>
<tr>
<td>2.</td>
<td>Description of the probes utilized in this study</td>
</tr>
<tr>
<td>3.</td>
<td>Sizes of the restriction fragments hybridizing to the alphoid family probe HH550-31</td>
</tr>
<tr>
<td>4.</td>
<td>Sizes of the restriction fragments hybridizing to the alphoid family probe HE340-9</td>
</tr>
<tr>
<td>5.</td>
<td>Sizes of the restriction fragments hybridizing to the β-satellite family probe H7-1</td>
</tr>
<tr>
<td>6.</td>
<td>Sizes of the restriction fragments hybridizing to the satellite III probe Kpn1.8</td>
</tr>
<tr>
<td>7.</td>
<td>Double digests of HH550-31 alphoid sequences in 153E7BX DNA</td>
</tr>
<tr>
<td>8.</td>
<td>Total DNA accounted for by the five probes utilized in this study</td>
</tr>
</tbody>
</table>
Relatively little is known about the mechanisms by which nondisjunction and translocation events occur at the chromosomal and sub-chromosomal level. Nondisjunction is the failure of chromosomes to segregate properly during mitosis or meiosis, and translocations involve the breakage and rejoining of fragments of chromosomes to other chromosomes. In humans, the acrocentric chromosomes (13, 14, 15, 21, and 22) participate more frequently than other chromosomes in these events (Therman, 1986), and would thus serve as a useful model to study the molecular mechanisms responsible for nondisjunction and translocation (Therman et al., 1989; Choo, 1990; Willard, 1990). In fact, this set of chromosomes account for 40% of all trisomies caused by nondisjunction (Hassold and Jacobs, 1984). In addition, the most common type of translocation, Robertsonian translocations, involve recombination between acrocentric chromosome short arms, with the 14-21 translocation being the most frequent (Thuline and Pueschel, 1982; Rosenmann et al., 1985).
Chromosome 21 is particularly important to study because of its frequent and well-documented involvement in the nondisjunction- and translocation-associated disorders such as Down Syndrome, Alzheimer Disease, premature centromere division, and cancers (Harris et al., 1986; Cooper and Hall, 1988). In each of these conditions, abnormalities involving chromosome 21 are thought to play a role in the onset of the condition. The short arm/centromere of chromosome 21 is a heterochromatic region consisting primarily of repetitive sequence DNA. The only actively transcribed genes are the ribosomal genes (Worton et al., 1984). The repetitive sequences present in this region have been specifically implicated in the nondisjunctions and translocations which are associated with the conditions mentioned above (Cooper and Hall, 1988). Therefore, determining the normal organization of the repetitive DNA on the short arm/centromeric region is essential to ultimately understanding its role in these events. In addition, this map would be useful in reconciling the physical and genetic maps of chromosome 21 and could also facilitate the use of restriction fragment length polymorphisms in determining the parental origin of a trisomy. Eventually, it could also be useful in detecting structural polymorphisms which predispose a chromosome to nondisjunction or translocation.

Previous work in our lab focused on identifying sequences on the short arm/centromere of human chromosome 21,
specifically on identifying the alphoid sequences present (Palamidis-Bourtsos, 1989; Chawla-Gupta, 1991). The specific purpose of this project was to construct a physical map for the entire short arm/centromeric region of chromosome 21 using pulsed field gel electrophoresis (PFGE) and hamster-human hybrid cell lines containing chromosome 21 or fragments of it as the sole human component. Five probes representing different tandem repetitive sequence families were utilized in this study: pHH550-31 and pHE340-9 (alphoid family probes), Kpn1.8 (a satellite III sequence), H7-1 (a β-satellite sequence), and CHB (a ribosomal gene probe).

Non-overlapping restriction fragments containing the repetitive sequences studied, including the calculated estimates of the ribosomal genes (Worton et al., 1988), total 21.2 to 31.9 Mb. This probably accounts for nearly all of the short arm/centromere region of chromosome 21. The alphoid families were found not to be significantly intermixed with each other, being separated by approximately 2 megabases (Mb). The HH550-31 alphoid family exists in at least two distinct clusters, one of which is in the centromeric region. The HE340-9 alphoid family cluster(s) are not interrupted by other sequences. β-satellite and satellite III sequences exist as multiple clusters on the short arm/centromere of the chromosome, and these two sequences are linked to each other at multiple sites both proximal and distal to the rDNA. The
\( \beta \)-satellite sequences are all distal to the centromeric region, while one satellite III cluster is centromeric.

The sequence organization found in the centromeric region of chromosome 21 is similar to that found in the centromeres of other human chromosomes (Manuelidis, 1978; Cooper et al., 1992 and 1993; Grady et al., 1992; Gravholt et al., 1992). This identifies structural features possibly required for centromere function. The multiple cluster arrangement of the short arm sequences observed in this study would facilitate frequent misalignment and out-of-register recombination between nonhomologous acrocentric chromosomes (Smith, 1976). This could explain the common occurrence of translocation and nondisjunction events involving these chromosomes.
Over the last three decades, extensive efforts have been made to understand the molecular mechanisms that underlie nondisjunction and translocation events. Nondisjunction is the failure of chromosomes to segregate properly during mitosis or meiosis, and translocations involve the breakage and rejoining of fragments of chromosomes to other chromosomes. At this point, relatively little is known about the mechanisms by which these events occur at the chromosomal and sub-chromosomal level. Studies of human nondisjunctions and translocations in malignancies, spontaneous abortions and live births revealed that these types of events are nonrandom; certain chromosomes are more frequently involved than others (Hassold and Jacobs, 1984; Heim and Mitelman, 1989; Willard, 1990). For example, the acrocentric chromosomes (13, 14, 15, 21, and 22) participate more frequently than other chromosomes in these events (Therman, 1986). In fact, this set of chromosomes account for 40% of all trisomies caused by nondisjunction (Hassold and Jacobs, 1984). In addition, the most common type of translocation, Robertsonian
translocations, involve recombination between acrocentric chromosome short arms, with the 14-21 translocation being the most frequent (Thuline and Pueschel, 1982; Rosenmann et al., 1985). The occurrence of 14-21 translocations taken together with the occurrence of 13-14 translocations account for 65% of all observed translocation events (Therman et al., 1989). Thus, the acrocentric chromosomes would serve as a useful model to study the molecular mechanisms responsible for nondisjunction and translocation (Therman et al., 1989; Choo, 1990; Willard, 1990). Chromosome 21 is particularly important to study because of its frequent and well-documented involvement in the nondisjunction- and translocation-associated conditions to be described below.

Down Syndrome

In 1866 Down syndrome (DS) was first identified as a developmental disorder by John Langdon Down. It is the most common genetic cause of retardation, occurring at a frequency of 1/600 live births (Cooper and Hall, 1988; Hassold and Jacobs, 1984). In addition to mental retardation, individuals with DS may exhibit such symptoms as congenital malformations of the heart, growth retardation, and hypotonia (Cooper and Hall, 1988). They also may have an increased risk of developing leukemia, hyperuricemia, and pathology similar to Alzheimer's disease, and usually have the characteristic facies and dermatoglyphic patterns (Cooper and Hall, 1988).
In 1959, it was determined that DS is caused by the presence of an extra copy of chromosome 21 (Lujeune et al., 1959). Trisomy 21 is the most common autosomal trisomy in live births (Hassold and Jacobs, 1984). Little is known of how this additional copy of chromosome 21 exerts its effects, but it is probably due to the combined effects of multiple genes on the long arm of the chromosome (Cooper and Hall, 1988). In fact, only a portion of the long arm, the q22.3 region, is sufficient when present in triplicate to induce a majority of the phenotypic characteristics of the condition (Carrit and Litt, 1989). Trisomy accounts for 95% of the known cases of DS, and Robertsonian translocations and mosaicism account for the remaining cases (Thuline and Pueschel, 1982). In recent molecular studies, Antonarakis et al. (1992) found that of 200 cases of trisomy 21 studied, 94% were maternal in origin, 4.5% were paternal, and in 1.5% of the cases, the origin could not be determined with the available markers. Of the trisomies of maternal origin, for which a stage of meiosis could be determined, 77.1% were meiosis I errors and 22.9% were meiosis II errors; the meiotic stage could not be determined for 11.7% of the maternal origin cases (Antonarakis et al., 1992). For the paternal origin trisomies, 22.2% cases were meiosis I errors and 77.8% were meiosis II errors (Antonarakis et al., 1992).

Only a few risk factors for DS have been identified. The most well-documented is advanced maternal age, with the
highest incidence in women over age 35 (Cooper and Hall, 1988). One familial study found that women under the age of 25 who had had a DS child had an increased risk of having another DS child (Lippman and Ayme, 1984). Another study found that a four-fold increase in the risk of having a DS child is present when the parents are closely related as compared to parents who are unrelated (Alfi et al., 1980). However, siblings and other relatives of DS individuals do not appear to be at an increased risk of having a DS child (Abuelo et al., 1986). There has been some suggestion that certain copies of chromosome 21 have an increased risk of meiotic nondisjunction (Hansson, 1979). A reduced rate of recombination for some copies of chromosome 21 results from the failure of normal chromosome pairing, predisposing the chromosome to nondisjunction (Warren et al., 1987). Thus, there is some suggestion that a genetic predisposition to nondisjunction may exist (Cooper and Hall, 1988), and the structure of the repetitive sequences on the short arm of chromosome 21 have been implicated in such events (Cooper and Hall, 1988). For example, a variation in the DNA sequences in the nucleolar organizing region (NOR), known as double NOR (dNOR), was found to be a risk factor for DS (Jackson-Cook et al., 1985; Verma et al., 1986).
Alzheimer Disease

Alzheimer Disease (AD) is a neurodegenerative disorder resulting in progressive and degenerative changes in memory, language, behavior, and cognitive function. It affects between 1.5 to 2 million people over 65 years of age in the United States alone, and it has been projected that by the year 2000, approximately 10% of the population over the age of 65 will be affected (Pettegrew, 1989). It is pathologically defined by brain lesions such as neuritic plaques (NP), neurofibrillary tangles (NFT), and neuronal loss (Selkoe, 1989). The NP contain a core consisting of a 42 amino acid protein, called amyloid, surrounded by degenerating neurons. This protein is a fragment of a 695 amino acid protein which is a normal constituent of brain tissue; the gene for this protein is located on the long arm of chromosome 21 (Katzman and Saitoh, 1991). NFT consist of a number of proteins such as tau, ubiquitin, and neurofilament protein and appear to be a characteristic of degenerating neurons rather than specifically AD, as they are present in a wide range of neurodegenerative disorders (Katzman and Saitoh, 1991).

The etiology of AD still remains somewhat of a mystery; the only risk factors that have been identified are age, family history and Down syndrome (Katzman and Saitoh, 1991). The risk of developing AD increases from 0.1% at age 65 to 47% at age 85 (Katzman and Saitoh, 1991). A genetic basis has been identified for only 15% of all cases, where an autosomal
dominant mutation is inherited in a form of AD termed Familial AD (FAD) (Katzman and Saitoh, 1991). FAD, which is an early onset form of the disease, has been associated with gene loci more proximal to the centromere than the amyloid gene on the long arm of chromosome 21 (Goate et al., 1989; Katzman and Saitoh, 1991). Not all forms of FAD are linked to chromosome 21 (Schellenberg et al., 1993; Van Broeckhoven et al., 1992). Families carrying the FAD gene have an increased risk of producing DS offspring (Katzman and Saitoh, 1991). Down syndrome appears to be the most important risk factor because all DS individuals over the age of 40 develop AD type pathology (Katzman and Saitoh, 1991). Because of these connections between DS and AD, Potter (1991) has proposed a unitary hypothesis to explain both the sporadic and familial forms of AD as well as its association with age and DS. He proposes that the accumulation of trisomy 21 somatic cells during adult life may underlie the sporadic and familial forms of AD. He suggests that the mutant gene in FAD may predispose the centromere to frequent mitotic or meiotic nondisjunction leading to an accumulation of trisomy 21 cells. As for the sporadic cases, environmental agents that promote improper chromosome segregation could result in the accumulation of trisomy 21 cells. In several cases, trisomy 21 cells have been isolated from AD patients (Potter, 1991). Thus, the hypothesis states that both AD and DS result from nondisjunction events leading to an accumulation of trisomy 21
cells with DS resulting from nondisjunction events in both the meiotic and mitotic processes and with AD resulting from only mitotic nondisjunction events. The symptoms are therefore the result of the number of trisomy 21 cells accumulated over time (Potter, 1991). Again, it should be noted that nondisjunction events leading to the formation of trisomy 21 cells are thought to result from changes in the organization or functional state of the repetitive DNA sequences on the short arm/centromere region of the chromosome (Cooper and Hall, 1988).

Premature Centromere Division

Premature centromere division (PCD) is an age-related functional defect of the centromere in which chromatids prematurely separate in the mitotic process resulting in nondisjunction (Fitzgerald et al., 1975). Several variations of the phenomenon have been observed, and it has been described in patients with DS, AD, Robert’s syndrome (RS), chronic myeloid leukemia, and as a dominantly inherited anomaly in three separate family studies (German, 1979; Tomkins, et al., 1979, Petrinelli et al., 1984; Moorhead and Heyman, 1983; Vig, 1984; Rudd et al., 1983). The familial forms of the phenomenon appear to involve all chromosomes separating prematurely at times. In one case study, chromosomes X, 18, and 21 were found to be the most frequently involved chromosomes (Fitzgerald et al., 1986). One group of
researchers has proposed that this phenomenon may result from out-of-phase separation of the centromeres of the chromosomes, since it has been shown that the separation sequence is a genetically controlled process (Vig, 1983; Vig and Wodnicki, 1974; Vig, 1981). Chromosome 18 is the first in the sequence and chromosome 21 is one of the last in the sequence, which may be a possible explanation for why these two chromosomes are more frequently affected by the defect (Mehes, 1978; Vig, 1981). In all cases, PCD results in repeated trisomies and has been found in parents of trisomy 21 offspring (Fitzgerald et al., 1986). It has also been suggested that PCD is the mechanism by which all human aneuploidies occur (Fitzgerald et al., 1975). The repetitive sequences on the short arm/centromere region are thought to play a role in centromere structure and/or function (Manuelidis, 1978); therefore, it is possible that these sequences are playing a role in the nondisjunction events in PCD. This emphasizes the necessity for understanding the organization of the repetitive sequences on the short arms/centromere region of the acrocentric chromosomes.

Cancer

In 1914, Theodor Boveri proposed what became known as the somatic mutation theory of cancer. His hypothesis suggested that nuclear changes, specifically chromosomal aberrations, were the cause of the transition of a cell from a normal to
neoplastic proliferative state. However, it was not until 1960 that the first neoplastic-associated aberration was identified in a patient with chronic myeloid leukemia: the Philadelphia chromosome (Nowell and Hungerford, 1960). The Philadelphia chromosome consists of a translocation of a portion of chromosome 9, around the abl oncogene, to chromosome 22 causing activation of the oncogene (Cooper, 1992). It has become evident, through cytogenetic and molecular studies, that, at the cellular level, cancer is a genetic disease (Heim and Mitelman, 1989). The genetic anomalies associated with cancers have been classified into three categories: primary and secondary abnormalities and cytogenetic noise. Primary abnormalities are those chromosomal aberrations that are essential in establishing the neoplasia. Secondary abnormalities result from the instability of the genome created by the tumor environment and are important in the progression of the tumor. Cytogenetic noise consists of chromosomal aberrations that occur within a tumor cell population and are temporarily detectable but are not evolutionarily beneficial to the tumor (reviewed in Heim and Mitelman, 1989). The chromosomal changes observed include monosomies and trisomies, translocations, inversions, and deletions (Heim and Mitelman, 1989). It is evident that the patterns of chromosomal changes in neoplasia are nonrandom (Heim and Mitelman, 1989), and in fact, often times, a specific chromosomal alteration is associated with a specific
tumorous condition (Rowley, 1973). For example, an 8;21
translocation is characteristic of acute myelogenous leukemia;
the breakpoint is in the q22.1 - q22.3 region of the
chromosome 21 (Rowley, 1973; Rowley, 1984). As a group, the
human acrocentric chromosomes are disproportionately involved
in nondisjunction and translocation events (Therman, 1986;
Jacobs et al., 1974; Therman et al., 1989) The nonrandomness
of the alterations are thought to be connected to the
molecular organization of the short arms of the acrocentric
chromosomes (Therman et al., 1989; Cheung et al., 1990).

The increased risk for cancer in the two most common
aneuploidies in man, DS and Klinefelter syndrome (47, XXY),
emphasizes the connection between aneuploidies and neoplasia
(Fong and Brodeur, 1987; Harnden et al., 1971). Chromosome 21
aneuploidy has been specifically associated with several
cancerous conditions. For example, trisomy 21 has been
proposed as the primary change initiating leukemogenesis
associated with DS (Sacchi, 1990) and has been associated with
transient myeloproliferative disorder (Hayashi et al., 1988).
Therefore, trisomy 21 is implicated in the proliferation of
megakaryoblastic cells (Bessho et al., 1988; Suda et al.,
1987). In addition, most studies show a 10-20 fold increase
in the risk of leukemia in DS patients relative to the general
population (Fong and Brodeur, 1987). Conversely, 20% of all
cases of acute leukemia involve a trisomy 21 aberration (Dube
and El-Solh, 1986; Fong and Brodeur, 1987). Furthermore, the
number, size, and location of the NOR has been associated with certain cancers (Crocker, 1990). In fact, silver staining of the NOR region can be used as a technique in distinguishing between benign and malignant tissue (Crocker, 1990). Translocations have also been noted to be nonrandom and often specific for a given tumorous condition (Rowley, 1973). In fact, in acute nonlymphocytic leukemias, chronic myeloproliferative disorders, acute lymphoblastic leukemia, chronic lymphoproliferative disorders, lymphomas, and solid tumors, translocations are the primary karyotypic aberration (Heim and Mitelman, 1989). These nondisjunction and translocation events involving the acrocentric chromosomes are thought to be dependent on the nature, amount, and organization of the DNA on the short arms (Cooper and Hall, 1988).

Acrocentric Chromosomes

The acrocentric chromosomes are the five human chromosomes carrying the ribosomal RNA genes (rDNA) on their short arms. The short arms are divided into three regions: the centromere-proximal p11 region, the rDNA p12 region, and the distal-telomeric p13 region, which are mainly heterochromatic regions with the only actively transcribed genes being the ribosomal genes. Despite the structural similarities of the acrocentric chromosomes' short arms and the fact that they are all composed of satellite and other
repetitive DNA, the nature and organization of these DNA sequences is highly variable among the acrocentrics (Choo, 1990; Chawla-Gupta, 1991). The region occupied by the rDNA is part of the NOR which participates in forming the nucleolus during most of the cell cycle (Cooper and Hall, 1988). During this interaction, formations known as satellite associations (SA) can occur (de Copoa, et al., 1973). SA are visible cross bridges of DNA between nonhomologous acrocentric chromosome short arms. An increased risk of meiotic nondisjunction appears to be associated with the predisposition of a chromosome to form satellite associations (Hansson and Mikkelsen, 1978; Hansson, 1979). The formation and strength of these associations is dependent on the amount and organization of the repetitive DNA present on the short arms and is heritable (Miller et al., 1977; Mattei, et al., 1976). In addition, the short arms of the acrocentric chromosomes interact physically with the X and Y chromosomes in mitosis and meiosis, and this association has also been implicated in translocation and nondisjunction events (Stahl et al., 1984; Mirre et al., 1980). Thus, it is evident that understanding the structure and organization of the sequences on the short arms/centromeres of the acrocentric chromosomes will facilitate our understanding of these events.
Chromosome 21

The involvement of chromosome 21 in the several conditions described above, in addition to several features of the chromosome itself, make it evident that this chromosome can serve as a particularly valuable model to study nondisjunction as well as translocation events (Harris et al., 1986; Cooper and Hall, 1988). Chromosome 21 is the smallest of all the autosomes, thought to contain 50,000 to 60,000 kilobase pairs (kb) of DNA, comprising approximately 1.4% to 1.9% of the total genome (Harris et al., 1986; Watkins et al., 1987). It is estimated to have 1000 to 2000 genes (Cooke and McKay, 1978). The long arm contains the Down syndrome pathological segment of the chromosome (Cooper and Hall, 1988). This segment is less than 3 megabases (Mb) in the q22.3 region of the chromosome, and appears to be sufficient, when present in triplicate, to cause the characteristic facies, mental retardation, short stature, clinodactyly, and gapped toes of DS (reviewed in Carrit and Litt, 1989). However the specific gene (or genes) responsible for these phenotypes has not yet been defined. Recent work suggests that the genes for other aspects of the DS phenotype are spread over larger regions of the long arm (Korenberg, 1990). To date, approximately 30 genes have been localized to the long arm of chromosome 21; however, the biochemical nature is only known for approximately one half of these genes (Carrit and Litt, 1989). Based on the genes mapped thus far,
there appears to be a gene rich area in the distal one third of the long arm while the proximal two thirds of the long arm appears to be gene poor (Gardiner, 1990). A physical map of the entire long arm (21q) has now been constructed using yeast artificial chromosome (YAC) (Chumakov et al., 1992). This map is in agreement with previous genetic linkage maps, PFGE restriction maps, and radiation induced hybrid maps; the majority of the markers showed the same relative positions on all maps (Chumakov et al., 1992). Having this map of the contigs will aid in the identification of the causes of the disorders mentioned previously (Chumakov et al., 1992). The YAC contig map as well as the recent NotI physical map of 21q suggests that this region of the chromosome is in the size range of 40 to 50 Mb, which is close to the size originally estimated for the entire chromosome (50 to 60 Mb) (Chumakov et al., 1992; Gardiner, 1990; Ichikawa et al., 1992).

The short arm is thought to contain 20,000 to 25,000 kb of DNA (Gardiner et al., 1988; Korenberg and Engels, 1978) consisting of a proximal stalk and a terminal knob or satellite. The stalk is divided into the proximal or centromeric (p11) region and the ribosomal DNA (rDNA) (p12) region, and the knob consists of the distal or telomeric (p13) region (Evans et al., 1974). The precise copy number and organization of ribosomal RNA genes on the short arm of chromosome 21 is not known. However, it is estimated that approximately 30-40 copies of the 44 kb repeat are present on
each of the acrocentric chromosomes (Worton et al., 1984). The remaining sequences on the short arm are repetitive in nature. It is changes in the organization or functional state of these DNA sequences that have been associated with DS, AD, PCD, and cancers (Choo et al., 1991; Cooper and Hall, 1988). While the long arm contains the genes responsible for the resulting phenotypes of the conditions, it is the structure and organization of the repetitive DNA sequences on the short arm that are the basic factors in causing the nondisjunction or translocation that is the initial cause of the conditions.

**Repetitive DNA**

Repetitive DNA makes up at least 40% of the human genome (Britten and Kohne, 1968). It is organized as either tandem repeats or as interspersed sequences. Interspersed sequences exist as single copies at multiple loci in the genome. Long interspersed repeated sequences (LINES) and short interspersed repeated sequences (SINES) are the two major types that exist in the human genome. Because of their structural properties, such as short direct terminal repeats, and their mobilities, both LINES and SINES sequences are thought to be transposons (Singer, 1982; Fanning and Singer, 1987). SINES sequences occur about $10^5$ times within the genome and are approximately 500 basepairs in length (Singer, 1982). Alu sequences are the best characterized example of a SINES sequence and are the most predominant repeat family in human DNA, making up 3 to 6%
of the human genome (Yoshinori et al., 1987). The average size of Alu sequences is 300 basepairs (bp), and there are approximately 300,000 to 500,000 copies in the haploid genome (Yoshinori et al., 1987). LINES sequences average 6000 bp in length (Adams et al., 1980). In the human genome, there are approximately $10^4$ copies of the LINES 1 (L1) family which is a heterogenous sequence (Singer and Skowronski, 1985). It has characteristic 3' A-rich stretches and truncations at the 5' end, and it also contains a long open reading frame (ORF) with homology to reverse transcriptase of viral and transposon origin (Singer and Skowronski, 1985; Hattori et al., 1986). Transcripts from L1 sequences have been detected in various cell types suggesting that they are active transposons (Hattori et al., 1986). The 724 family is a low copy number interspersed repetitive DNA located throughout the pericentromeric region of the acrocentric chromosomes, but not within the rDNA clusters (Kurnit et al., 1986).

Tandem repeats have a basic unit of repetition that exists as adjacent multiple copies. They are classified as either simple, with a repeating unit of less than 10 bp, or complex, with a repeating unit of greater than 10 bp (Gray et al., 1985). The satellite families, I, II, III, and IV, are one example of simple tandem repetitive DNAs. They are made up of short oligonucleotide tandem repeat units organized in long chromosome-specific arrays (Waye and Willard, 1989). The satellite DNA families all have similar patterns of
hybridization to centromeric heterochromatin (Higgins et al., 1985). $\beta$-satellite is a complex tandem repetitive DNA localized to a number of regions in the genome including the short arms of the acrocentric chromosomes (Waye and Willard, 1989). It has a 68-69 bp repeating unit and can be divided into at least two distinct subsets, one of which is shared by the acrocentrics (Waye and Willard, 1989).

The most abundant and best-characterized complex tandemly repeated DNA in humans is alphoid DNA. Alphoid sequences are classified into families and subfamilies according to their degree of sequence relatedness (Willard and Waye, 1987). Presently eleven different alphoid families have been identified (Willard and Waye, 1987a; Waye and Willard, 1987; Vissel and Choo, 1987). Chromosome specific alphoid families have been identified, and some studies indicate that each chromosome may have a specific organization of alphoid DNA (Willard and Waye, 1987a). All families consist of a basic repeating unit of about 170 bp arranged in long tandem clusters (Willard and Waye, 1987a). There are approximately 500,000 copies of alphoid DNA in the genome (Shmookler-Reis and Goldstein, 1980) which are primarily present at the centromeric/pericentromeric regions of chromosomes (Manuelidis, 1978). Because of this location and the sequence similarity to the known functional centromeric sequences in yeast (Fitzgerald-Hayes et al., 1982), alphoid sequences are thought to play a role in centromere structure or function.
(Manuelidis, 1978). In addition, the human centromere antigen, CENP-B, specifically binds to a region in some alphoid sequences (Masumoto et al., 1989).

Repetitive Sequences Localized on Chromosome 21

Several alphoid as well as several non-alphoid repetitive sequences have been found to be present on chromosome 21. Five different alphoid families, each with a unique organization and copy number, are present on chromosome 21, and it is known that these families are all located on the short arm of the chromosome (Palamidis-Bourtsos, 1989). It has also been determined that these alphoid sequences are organized as multiple clusters (Chawla-Gupta, 1991). The HH550 and HE340 families are discussed below. p308 is an alphoid family that is located at the centromeres of all autosomes and the X chromosome although the organization varies from one chromosome to another (Jabs et al., 1984). The pTRA2 alphoid sequence was first isolated from a chromosome 21 library, and, while it is found on all of the chromosomes, is particularly abundant on the acrocentric chromosomes (Choo et al., 1988). The BamHI 2.0 kb family was thought to be specific to the X chromosome (Yang et al., 1982; Willard et al., 1983), but copies of the sequence were also found on chromosome 1 (Carine et al., 1989) and on chromosome 21 (Palamidis-Bourtsos, 1989). On chromosome 21, this family
exists as 1.36 kb AluI and EcoRI repeats with 50 copies on the chromosome (Palamidis-Bourtsos, 1989).

The tandemly repeated ribosomal RNA genes, the satellite III family, and the \( \beta \)-satellite family have also been localized to the short arm of chromosome 21 (Watkins et al., 1987). \( \beta \)-satellite is a tandemly repeated DNA consisting of an array of diverged 68 bp monomer repeat units (Waye and Willard, 1989). Two distinct subsets have been localized proximal (between the centromere and rDNA) and distal to the rDNA on chromosome 21 (Willard, 1990). The satellite III subfamily localized to the short arm of chromosome 21 has a five bp repeat and has been localized, like \( \beta \)-satellite, proximal and distal to the rDNA (Doering et al., unpublished observation).

In addition to the tandem repetitive sequence families localized on the short arm, there are also a few interspersed repetitive sequences: Alu, L1, and 724. 724 is a low-copy-number interspersed repetitive sequence family in the pericentromeric regions of the acrocentric chromosomes (Kurnit et al., 1986). It is known not to be within the rDNA clusters (Kurnit et al., 1986). The Alu SINES family is underrepresented on the short arm of chromosome 21 as compared to the entire genome (Sainz et al., 1992). The L1 LINES family has a different distribution on the short arm as compared to the long arm, and the copies that are present appear to be full length, whereas in other locations the 5' ends are
truncated (Doering et al., 1988b). These studies indicate that the organization of repetitive DNA on chromosome 21 is very complex.

**DNA Sequence Probes Utilized in this Study**

Several of the families of repetitive DNA that have been localized to the short arm of chromosome 21 will be used in this study to determine a physical map for this region. These include two alphoid families as well as several non-alphoid families. These sequences were chosen because they are the major families present on the short arm of the chromosome. A brief description of each of the families is presented below.

1. **340 bp Eco RI family**: This alphoid family consists of a basic repeating unit of two tandem subunits, 171 bp and 169 bp long (Wu and Manuelidis, 1980), and it is present in approximately 175,000 copies in the haploid genome, 2-3% of the total genome (Shmookler-Reis and Goldstein, 1980; Manuelidis, 1978). This sequence is found, in particular abundance, at the centromeres of chromosomes 1, 3, 7, 10, and 19 as well as at the centromeres of all the acrocentrics (Manuelidis, 1978). There are several variants of this family which are classified into subfamilies based on their sequence relatedness (Willard, 1985; Jorgensen et al., 1986; Jorgensen et al., 1987; Willard and Waye, 1987a). Three subfamilies of the 340 bp family have been found on chromosome 21, and each subfamily has its own organization pattern in the genome and
on chromosome 21 (Palamidis-Bourtsos, 1989). The 340 family on chromosome 21 is organized as a continuous block with no unrelated sequences interrupting it, occupying approximately 2,500-3,000 kb of DNA (Doering et al., 1988a; Palamidis-Bourtsos, 1989; Chawla-Gupta, 1991). It has been localized to the centromeric region proximal to the rDNA of the chromosome (Palamidis-Bourtsos, 1989).

2. **pHH550 family**: This recently isolated 550 bp HindIII alphoid family is unusually heterogenous in sequence (Doering et al., 1988a). It is found on chromosome 21 as well as other chromosomes and exhibits chromosome specific organization (Palamidis-Bourtsos, 1989). It is organized predominately as 1.1 kb AluI and EcoRI fragments on chromosome 21 (Palamidis-Bourtsos, 1989). This family, unlike other alphoids, has a high degree of similarity to sequences found in all primates tested, including monkeys, which indicates it may be an evolutionarily old sequence (Carnahan et al., 1993). In the haploid genome there are approximately 44,000 copies, and approximately 2,500 of these copies are present on chromosome 21. The pHH550 sequences occupy around 1350 - 1400 kb of DNA (Palamidis-Bourtsos, 1989), but unlike the 340 family, this family exists in at least two distinct clusters with some unrelated sequences interspersed within the clusters (Chawla-Gupta, 1991). It does map to the same general region of the chromosome as the 340 family (Palamidis-Bourtsos, 1989), but the detailed organization of these two families relative to
each other is not known. It is known that these two families are separated by at least 2 million base pairs (Chawla-Gupta et al., 1991).

3. **H7 family**: This tandem repetitive sequence family is found only on chromosome 1 and the five acrocentric chromosomes (Devine et al., 1983). It appears to be present in low to moderate copy number in the human genome and is organized as clusters at the centromeres and/or short arms of the acrocentric chromosomes (Graham et al., 1985). There has been some evidence that these sequences show linkage with the rDNA and the 724 family (Kurnit et al., 1986). However, they appear to be rarely interspersed with KpnI (LINES 1) or Alu sequences and do not appear to be linked to alphoid sequences (Graham et al., 1985). The probe used in this study is H7-1 which is a 2.9 kb fragment subcloned from a larger genomic clone (Graham et al., 1985). There is 70% sequence similarity of this probe to β-satellite, and so it appears that H7 is a subfamily of β-satellite (unpublished observation, McCutcheon, et al.). β-satellite sequences have been found to be interspersed with two different additional repetitive sequences (Meneveri et al., 1993; McCutcheon et al., unpublished observation).

4. **Satellite III family**: The satellite III repetitive sequence family is a tandemly repeated pentamer with a consensus sequence of 5'–AATGG–3' (Higgins et al., 1985). The probe utilized in this study, Kpn1.8, is a 1.8 kb higher order
repeat defined by KpnI sites that originally appeared to be specific to the p11 region of chromosome 15 (Higgins et al., 1985). There are 3,000 copies of this repeat in the diploid human genome which are organized in long tandem arrays of perfect and imperfect repeats of the underlying pentamer consensus (Higgins et al., 1985). Sequences related to the Kpn1.8 subfamily have been found, thus far, on all the acrocentric chromosomes, where they appear to have a clustered organization (Doering et al., unpublished observation). They have a distinct pattern of organization on chromosome 21 as compared to the total genome (Doering et al., 1988b). On chromosome 21, these sequences are specific to the short arm (Doering et al., 1988b) and appear to be on both the proximal and distal side of the rRNA genes (Doering et al., unpublished data).

5. **Human ribosomal RNA genes**: Human ribosomal RNA genes (rDNA) are found in the middle of the short arms of the acrocentric chromosomes. There are 300 to 400 copies in the human genome, and they exist as tandem repeat clusters and are transcribed in a telomere to centromere direction (Worton, et al., 1988). The rDNA block lies within the NOR (Goodpasture and Bloom, 1975) and consists of a 44 kb repeating unit which includes 31 kb of nontranscribed spacer (Wellauer and Dawid, 1979; Sylvester et al., 1986). The spacer contains a region that varies in length from one repeat to the next (Worton et al., 1988). The probe that will be used in the present study,
$C_{HB}$, is a constant portion of the spacer region since only the spacer is a species-specific sequence and thus will not cross-hybridize with the hamster DNA in the hybrid panels (Worton et al., 1988). $C_{HB}$ is a 400 bp BamHI-HindIII fragment from a portion of the spacer, termed the C region of the repeat (Sylvester, et al., 1986).
MATERIALS AND METHODS

Experimental Approach

Several steps are required in constructing a physical map of the short arm of human chromosome 21. The first step was to establish linkage between different families of repetitive sequences. This was accomplished using pulsed field gel electrophoresis (PFGE), because traditional gel electrophoresis techniques cannot resolve the large DNA fragments (greater than 50 kilobasepairs) required to map long regions of a chromosome (Tyler-Smith and Brown, 1987). PFGE, specifically the Beckman Transverse Alternating Field Electrophoresis (TAFE) system, can resolve fragments ranging from 0.2 to 7.0 Megabases (Mb) using various protocols (Avdalovic and Furst, 1988; Wert and Furst, 1988). Linkage between families was determined by hybridizing the membrane-bound DNA with one probe, stripping the membrane, then rehybridizing the same membrane with other probes. Hybridization stringencies were such that there was no cross hybridization between the probes used in this study. The
autoradiograms obtained from each of the different probes were then superimposed, allowing an accurate determination of whether or not two probes were indeed hybridizing to the same fragment (Gardiner et al., 1988; Palamidis-Bourtsos, 1989; Chawla-Gupta, 1991). Two families were considered to be physically linked if their probes hybridized to the same size pulsed field gel fragment, and such fragments were produced by at least two different enzymes (Gardiner et al., 1988).

Once linkage between probes was established, hybrid cell lines containing different pieces of chromosome 21 were used to determine the general position of each fragment on the short arm. Depending on where the family is located, the probe will hybridize to the DNAs of some of the cell lines and not to others.

Double digests were used to further establish the organization of the HH550-31 alphoid family on the 153E7BX restriction fragments. Double digests were done by either sequential or simultaneous digestions of the DNA with two enzymes before running the gel. Using the data obtained from the above procedures, methods analogous to standard restriction mapping methods were used to construct a more detailed physical map of the short arm of chromosome 21 (Winberg and Hammarskjold, 1980; Warburton and Willard, 1990).
DNA Sources

The cell lines that were used are Chinese hamster-human hybrid cell lines in which chromosome 21 or fragments of it are the sole human component (received from Dr. M. Cummings of the University of Illinois, Chicago). Cell line 153-E9A contains the whole chromosome 21 (Moore et al., 1977); 153-E7BX lacks the short arm, and 2Fu'l lacks the short arm and the centromere (VanKeuren et al., 1986). The cell lines were cultured in 25 cm² flasks in approximately 10 ml Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum, 50 U/ml penicillin, and 50 mcg/ml streptomycin and maintained at 37°C and 5% CO₂.

Preparation of DNA samples

Cell cultures were washed with approximately 3 ml of 1xPBS (Phosphate Buffered Saline, 10x stock contains: 1.0 M NaCl, 0.4 M Na₂HPO₄, and 0.3 M NaOH, pH 6.9 to 7.4) and trypsinized in approximately 2 ml of a 0.625% trypsin (BRL, 2.5% stock solution was diluted in 1 x PBS) solution. The reaction was stopped with approximately 3 ml of complete medium. The cells were then pelleted by centrifuge action, resuspended in approximately 1 ml of complete medium, pooled into a single tube and recentrifuged to form one pellet. These harvested cells were resuspended to a concentration of 8-10 x 10⁶ cells/ml in 1xPBS. This solution was maintained at 50°C. An equal volume of 2% low melting point (LMP)
agarose (Beckman) was added to the cell suspension and mixed thoroughly. This mixture was then pipetted into a gel mold (Beckman) and left on ice at 4°C for approximately 2 hours to gel. These agarose plugs were then removed from the plug mold and washed at 50°C, shaking gently for 1-3 hours in approximately 5 volumes of ES (0.5 M EDTA, pH 9.0 with 1% N-lauroyl sarcosyl) containing 1 mM p-chloromercuri-benzoic acid (pCMB), which is a nuclease inhibitor added to help prevent the DNA degradation that can occur during these procedures. The dimension of the plugs is 0.25 x 0.25 x 2.5 cm; therefore, the volume is 160 ul and 5 volumes would be 800 ul per plug (approximately 1 ml per plug was usually used). The plugs were then incubated for four days at 50°C with gentle shaking in approximately 5 volumes ES with 1 mM pCMB and 2 mg/ml of proteinase K, changing the solution once every 24 hours, to lyse the cells and release the DNA into the agarose. Next, the plugs were rinsed in 5 volumes ES with 1 mM pCMB at room temperature two times for 15 minutes each. Four washes were then done for 1-2 hours each, agitating gently on a rotator in 20-25 volumes of Tris-EDTA, pH 7.4 (10 mM Tris, 1 mM EDTA) with 1 mM pCMB and 1 mM phenylmethylsulfonylfluoride. The plugs were then rinsed with several changes of distilled water at room temperature and then washed at least 3 times over a 2-3 hour period agitating gently on a rotator at 4°C in 20-25 volumes of Tris-EDTA, pH 7.4 with 1 mM pCMB. One final overnight wash was
done at 4°C agitating gently on a rotator in 20-25 volumes of T₅E₀.₅, pH 7.4 (5 mM Tris, 0.5 mM EDTA) with 1 mM pCMB. The plugs were then stored at 4°C in T₅E₀.₅, pH 7.4 with 1 mM pCMB until used (Jabs et al., 1989).

**Restriction Enzyme Digestions**

Restriction enzymes that cut infrequently in the genome were used to generate large fragments of DNA for analysis by PFGE. Restriction digests were performed by incubating the plugs at 37°C for 15 hours in 5 volumes of the appropriate restriction buffer containing 50 μg/ml BSA, 7 mM β-mercaptoethanol, 1 mM pCMB, and 30-40 units of restriction enzyme per lane of DNA (0.8 μg of DNA). The dimensions of the DNA plug for one gel lane are 0.25 x 0.25 x 0.50 cm; therefore, the volume of each lane of DNA was 32 μl. The digestion reaction was stopped by first rinsing each plug two times for 15 minutes each at room temperature in 1 ml of T₁₀E₁, pH 8.4 containing 0.5% N-lauroyl sarkosyl and 1 mM pCMB. They were then rinsed 3 times for 20-25 minutes each at room temperature in 1 ml of T₁₀E₁, pH 8.4 containing 1 mM pCMB to equilibrate them before running the gel (Jabs et al., 1989; Gardiner et al., 1988). Double digests were performed as either sequential or simultaneous digests. Sequential digests were performed by digesting the plug in the first enzyme as usual. However, after the standard 15 hour incubation, the
plug was then rinsed with only $T_{10}E_1$, pH 8.4 to remove the first enzyme. The second enzyme was then added in its own reaction buffer, and the digestion was allowed to proceed for another 15 hours. After this second incubation, the reaction was stopped as described above. Simultaneous digests were performed by adding both enzymes to the digest solution, using the reaction buffer with the highest salt concentration, incubating for the standard 15 hours and stopping the reaction as described above.

**TAFE Gel Conditions**

For a brief review of the TAFE gel programs used in this study, refer to table 1.

Three separate ramped protocols, each with an optimum range of resolution, were used to resolve the range of fragment sizes generated with the restriction enzymes used. The running buffer, TAFE buffer (10 mM Tris, 0.5 M EDTA, 43.5 mM acetic acid) precooled at 4°C for 1-2 hours or overnight before use, was used with each protocol.

The "low" gel runs were used to resolve fragments in the size range of 0.20 to approximately 1.00 Mb. The low gels were 1% LE (Low Endosmosis) agarose (Beckman) in TAFE buffer cooled to 4°C to mold. The DNA plugs were gently pushed into the gel wells using a sterile spatula and pipette tip. One percent LE agarose was then used to seal the wells prior to the run. The gel was then placed in the electrophoresis box
with the running buffer maintained at 15°C. The ramped protocol consisted of two stages. The first was 30 minutes long with a pulse time of 4 seconds and a current of 170 mA to drive the DNA into the gel. The second stage was run at 150 mA for 20 hours with a 60 second pulse (Avdalovic and Furst, 1988).

The "medium" gel runs were used to resolve fragments in the size range of approximately 0.70 to 2.00 Mb. These gels were 0.7% LE agarose in TAFE buffer. The samples were loaded as above. The wells were sealed using 0.7% LE agarose. The running buffer was maintained at 10°C for the duration of the run. The first stage of the medium run was the same as the first stage of the low gels. The second stage was 20 hours long with an 80 second pulse time at 150 mA.

The "high" gel runs were used to resolve fragments in the 2.00 to 7.00 Mb size range. A 0.6% LE agarose gel in TAFE buffer was used for these runs, and the running buffer was maintained at 10°C. Again, samples were loaded as described above, and the wells were sealed using the 0.6% LE agarose. The first stage ran for 3 hours at 40 volts with a 45 minute pulse time. The second stage ran for 87 hours with a pulse time of 30 minutes at 60 volts (Mary Morgenstern, personal communication).

All gels were stained in 200 mL of a 0.01% ethidium bromide solution in 1xE buffer (0.04 M Tris; 0.02 M NaAc; 0.002 M EDTA, pH 7.8) for 30-40 minutes and photographed under
ultraviolet light. \textit{S. cerevisiae} and \textit{S. pombe} markers (Beckman) were used to determine the size of the fragments on the gels (Smith et al., 1987; Chawla-Gupta, 1991). All marker plugs were equilibrated prior to use by washing 2 times for 30 minutes each in 20–25 volumes of $\mathrm{T_{10}E_{1}}$, pH 7.4 rotating gently at room temperature.

\textbf{Transfer and Hybridization}

All gels were blotted and hybridized in the same manner. The DNA was transferred to a nylon membrane (Gene Screen Plus hybridization transfer membranes; NEN Research Products) overnight using the alkali Southern blot method (Reed and Mann, 1985) or using the alkali method on a Vacugene XL vacuum blotting apparatus (Pharmacia). Prehybridization of the membranes proceeded for at least 5 hours at 37°C with agitation in either 20 ml or 40 ml of hybridization solution. Probes were labeled with $^{32}$P by Nick translation as described by Doering et al. (1982). Hybridization with the radioactively labeled probe (Doering et al., 1982) was allowed to proceed overnight at 37°C with agitation (Hybridization buffer: 50% formamide; 1 M NaCl; 50 mM Tris, pH 7.5; 1% sodium dodecy sulfate, SDS; 10ug/ml heat denatured \textit{E. coli} DNA; 6 ng/ml heat denatured probe). Different probes representing different sequences localized to the short arm of human chromosome 21 were used. Two probes representing two different alphoid sequence families, Eco RI 340 (Manuelidis,
1978) and pH550 (Doering et al., 1988a), were used as well as a probe representing the H7 family (Graham et al., 1984). A probe representing a satellite III sequence (Doering et al., 1988b) and one for the spacer region of the ribosomal RNA genes (Worton et al., 1988) were also used. Standard hybridization wash conditions using agitation were as follows: two times for 15 minutes each at room temperature in 50 ml of 2x SSC (1x SSC is 150 mM NaCl, 15 mM sodium citrate, 0.1 mM EDTA); two times for 30 minutes each at 60°C in 100 ml of 2x SSC with 1% SDS; and lastly, two times for 30 minutes each at room temperature in 50 or 100 ml of 0.5x SSC. After the washes, the membrane was sealed in a plastic bag (Brl) and autoradiographed using Kodak XAR film with or without Cronex intensifying screens.

After autoradiographs were obtained for one probe, the membrane was stripped of this probe by one of two methods. The first is the NaOH method which consists of two washes. The first is a 30 minute wash at 42°C with agitation in 50 ml of 0.4 M NaOH. The second wash is again 30 minutes at 42°C with agitation but in 50 ml of a solution of 0.1x SSC, 0.1% SDS, and 0.2 M Tris, pH 7.5. The second method is the boiling method. It consists of a 30 minute wash with the membrane submerged in a boiling water bath in 200 ml of a solution of 10 mM Tris, pH 7.5; 1 mM EDTA; and 1% SDS. Stripped membranes were autoradiographed for at least twenty-four hours with an
intensifying screen to ensure that no residual probe remained before the membrane was rehybridized with another probe.
RESULTS

Determination of Restriction Fragments

DNA from the hamster-human hybrid cell line 153E9A, which contains whole human chromosome 21 as the sole human component (Figure 5), was digested with various restriction enzymes that do not cut within the repetitive sequence families. Three different pulsed field gel electrophoresis programs, resolving fragments from 0.2 to 6.0 Mb (reviewed in Table 1), were used in this study. The gels were then blotted to nylon filters using the Southern alkali method and probed with radiolabeled repetitive sequences known to be on the short arm/centromere region of chromosome 21. For each probe, the restriction enzymes which generated fragments containing the largest portion of the short arm/centromere DNA were chosen for study. These restriction fragments were subsequently used in calculations to establish the most inclusive map of the short arm/centromere with the sequences utilized in this study. Enzymes were not considered if they cut the DNA of a sequence family into unresolvable fragments. Such fragments could either be too small, running off the gel or in the compacted
fraction at the bottom of the gel, or too large to enter the gel. Only fragments that demonstrated strong hybridization were considered since these contained most of the sequence being studied. Fragments which hybridized faintly were not considered since such fragments likely contained only a few copies of the sequence. Table 2 lists the probes and contains a brief description of each for easy reference.

**HH550-31:** For the alphoid probe HH550-31, the restriction enzymes which produced useful bands with the TAFE programs used were BamHI, Clal, SalI, and SfiI. Other enzymes that were tested included KpnI, PvuII, SstII, MluI, XhoI, NotI, BsshII, and SacII; however, these enzymes did not produce useful fragments with this probe. The four restriction enzymes used in this study each generated four to six fragments to which this probe hybridized, ranging in size from 0.29 Mb to 4.14 Mb (Table 3). In each of the restriction digests, HH550-31 hybridized to four fragments greater than one Mb. In three of the digests, there were additional fragments less than one Mb also present. This indicates that HH550-31 is likely arranged in several clusters, since if it were arranged in a single or few blocks on the chromosome, then only one or a few restriction fragments would be generated. This organization is supported by its localization data (discussed below), and previous data obtained by Chawla-Gupta (1991). Digestion with SfiI produced four large fragments, 4.14, 1.62, 1.56 and 1.51 Mb, totaling 8.83 Mb, the
largest amount of DNA accounted for by the HH550-31 probe. Copy number estimates for this sequence on the short arm/centromere region indicate that it occupies a total of only 1.35 to 1.4 Mb of DNA (Palamidis-Bourtsos, 1989; Chawla-Gupta, 1991). This amount of sequence is much less than the 8.83 Mb total of DNA fragments to which the HH550-31 probe hybridizes, supporting the existence of multiple clusters of the HH550-31 sequence on the chromosome with other unrelated sequences interspersed. Examples of HH550-31 hybridization to 153E9A DNA digests are shown in Figure 1, lanes 1, 3, 5, 7 and 8.

**HE340-9:** The HE340-9 probe when hybridized to digests of 153E9A DNA did not hybridize to as many fragments, nor as many large fragments (greater than one Mb), as did the HH550-31 probe. KpnI, SalI, SstII, MluI, XhoI, and NotI restriction enzymes were tested and found not to produce resolvable fragments with the pulsed field gel electrophoresis programs used in this study. The enzymes which did produce useful fragments were BamHI, ClaI, PvuII, and SfiI. The fragments produced ranged in size from 0.11 Mb (Chawla-Gupta, 1991) to 3.30 Mb. Table 4 lists the fragments produced when 153E9A DNA was digested with each of the enzymes and probed with HE340-9. Only two to four fragments were detected with each of the enzymes. This would suggest that HE340-9 exists as few, or even one, cluster(s) on the short arm of the chromosome, which is consistent with data obtained previously (Chawla-Gupta, 1991). Figure 2 illustrates the fragments produced with
various enzymes. The largest amount of DNA accounted for by this probe was generated with the restriction enzyme SfiI which produced two large bands, 2.70 and 3.26 Mb, totaling 5.96 Mb. Based on the number of copies of the HE340-9 sequence present on the short arm/centromere of chromosome 21, the sequence occupies 2.5 to 3.0 Mb of DNA (Palamidis-Bourtsos, 1989; Chawla-Gupta, 1991). The two fragments totaling 5.96 Mb most likely contain nearly all of the HE340-9 sequence, again suggesting that this family exists in one or a few clusters.

**H7-1:** Five restriction enzymes produced resolvable fragments of 153E9A DNA when the digests were hybridized with the H7-1 probe. They are ClaI, SalI, SfiI, SstII, and MluI. Enzymes that did not produce resolvable fragments included BamHI, PvuII, KpnI, XhoI, and NotI. The H7-1 probe hybridized to three to five fragments with each of the various enzymes. The fragments generated ranged in size from 0.36 Mb to 2.37 Mb. MluI produced six bands totaling 7.17 Mb, the largest amount of DNA accounted for by the H7-1 probe. The H7-1 sequence is thought to be present in low to moderate copy number (Graham et al., 1985); therefore, the large total obtained, 7.17 Mb, would suggest that this sequence exists in multiple clusters. The fragments generated with each of the enzymes are listed in Table 5. Examples of H7-1 hybridization to digests with these enzymes are shown in Figure 3.
Kpn1.8: Kpn1.8 detected four to eight fragments with the restriction enzymes SalI, SfiI, ClaI, SstII, and MluI. Other enzymes that were tested but did not produce resolvable fragments included XhoI, NotI, PvuII, BamHI, and KpnI. The large number of fragments to which the Kpn1.8 probe hybridized suggests that this sequence exists as multiple clusters. The restriction enzyme MluI generated six fragments to which Kpn1.8 hybridized, totaling 12.16 Mb. This large total would also indicate that satellite III exists as multiple clusters. Such a finding is consistent with the studies of Doering et al. (unpublished observation), which demonstrate the presence of clusters of satellite III sequences both proximal and distal to the rDNA on the short arm. Thus the 12.16 Mb total must represent at least two, most likely more, clusters of this sequence. Table 6 lists the fragments produced with each of the enzymes used, and Figure 4 (lanes 1, 2, 3, and 5) shows examples of the fragments produced by these enzymes.

Sequence Localization

Each sequence was localized to a specific region of chromosome 21 by hybridization to digests of DNA from hamster-human hybrid cell lines containing the whole human chromosome 21 or only portions of it as the sole human component (Figure 5). The 153E9A cell line contains the whole human chromosome 21, 153E7BX contains the long arm and centromeric region of chromosome 21, and 2Fu'1 contains only the long arm (Figure
None of the probes utilized in this study hybridized to the 2Fu'1 DNA, therefore demonstrating that all these probes hybridize to sequences on the short arm/centromeric region of the chromosome (not shown). The probes CHB, pHE340-9, and pH7-1 do not hybridize to digested 153E7BX DNA. This is demonstrated for pHE340-9 in lanes 3 and 5 of Figure 2, for pH7-1 in lanes 6 and 8 of Figure 3, and for CHB in lanes 2 and 4 of Figure 9. These sequences are therefore present on the short arm of the chromosome outside of the centromeric region.

The alphoid HH550-31 probe and the satellite III Kpn1.8 probe do hybridize to 153E7BX DNA in addition to 153E9A DNA, indicating that these two sequences are more centromeric in location than the other three sequences studied. The fragments present in the 153E7BX digests represent a subset of the fragments present in the 153E9A digests. If a sequence is present in a fragment that spans the breakpoint of the 153E7BX cell line, then one of the fragments present in the 153E7BX digest will not correspond in size to a fragment in the 153E9A digest, being either larger or smaller than the corresponding 153E9A fragment. A larger sized fragment is possible since the 153E7BX cell line fragment of chromosome 21 is a translocation to a portion of the hamster DNA (VanKeuren et al., 1986). If a sequence does not span the breakpoint, then all fragments present in the 153E7BX digests will be the same size as ones present in the 153E9A digests. If this is true then that sequence must exist as a distinct cluster in the
centromeric region, separate from clusters outside the region.

For the HH550-31 probe, all the fragments found in the 153E7BX digests are also present in the 153E9A digests (Figure 1, lanes 1-6) indicating that this sequence does not span the 153E7BX breakpoint and confirming the existence of multiple clusters of this sequence. For Kpn1.8, one fragment in the 153E7BX digest, with two different enzymes, is not found in the 153E9A digest. This indicates that this sequence spans the breakpoint of the 153E7BX portion of chromosome 21 (Figure 4, lanes 3-6). The 153E9A Sall digest produces six bands ranging in size from 1.29 Mb to 0.72 Mb. The 153E7BX Sall digest produces only a single band of 0.63 Mb which is smaller than any of the 153E9A fragments. The SfiI digests probed with Kpn1.8 also demonstrate the Kpn1.8 sequence spanning the 153E7BX breakpoint. Two bands hybridize to the Kpn1.8 probe in 153E7BX; one is the same size as a band in the 153E9A digest, 0.69 Mb, and the other is smaller than any of the fragments in the 153E9A digest at 0.50 Mb. Observing differences in fragment sizes between 153E9A and 153E7BX DNA digests with two different enzymes strengthens the conclusion that Kpn1.8 (satellite III) sequences span the 153E7BX breakpoint.

**Linkage**

To establish linkage between two sequences, probes for both had to hybridize to the same size restriction fragment
made by at least two different enzymes. Southern blots were stripped of one probe and rehybridized with a different probe as described in the Materials and Methods section. The stripped blots were exposed to film to eliminate the possibility of residual hybridization. To confirm linkage, fresh Southern blots, using the same restriction enzymes and TAFE programs as the first blot, were hybridized in the opposite probe order as the first blot. In this manner the hybridization of two probes to the same size fragment could be unequivocally established. Using these criteria, linkage was established between the two alphoid family probes, HH550-31 and HE340-9 (Figure 6), and between the β-satellite probe, H7-1, and the satellite III probe, Kpn1.8 (Figure 7).

HH550-31 and HE340-9 were found to be linked on a 3.3 Mb ClaI fragment (Figure 6, lanes 3 and 4) and on a 3.0 Mb BamHI fragment (Figure 6, lanes 1 and 2). Both linked fragments were confirmed by the above criteria (Tables 3 and 4). Linkage was also observed but not confirmed on a 2.7 Mb KpnI fragment (not shown), since KpnI did not produce resolvable fragments for the majority of the alphoid DNA sequences. These results suggest that clusters of the two alphoid sequences are separated by at least 2 Mb which, is consistent with previous data (Chawla-Gupta, 1991).

H7-1 and Kpn1.8 were found to be linked on multiple fragments of enzyme digests with several different enzymes. MluI and SalI digests of 153E9A DNA produced three fragments
each to which both H7-1 and Kpnl.8 hybridized. The SalI linked fragments were 1.18, 1.06, and 0.72 Mb in size, and the MluI linked fragments were 1.50, 1.36, and 1.26 Mb in size (Figure 7, lanes 1-4; Tables 5 and 6). Again, the linkage observed was confirmed by reversing the order of probe hybridization to a fresh Southern blot. The restriction enzymes ClaI and SfiI produced one linked fragment each. The ClaI fragment is 0.64 Mb (not shown), and the SfiI fragment is 0.56 Mb (Figure 7, lanes 5 & 6). The presence of three linked fragments produced by at least two different enzymes supports the distribution of the H7-1 and Kpnl.8 sequences in multiple clusters because if only one cluster of each existed, then only one linked fragment could exist for each enzyme digest.

No linkage was established between either of the alphoid probes and H7-1 or Kpnl.8. This is evident since these sequences do not share any common fragments (Tables 3, 4, 5 and 6). Autoradiograms demonstrating this are presented in Figure 8. Lanes 1 through 3 are 153E9A DNA BamHI digests hybridized with the H7-1, HH550-31, and Kpnl.8 probes respectively. Lanes 4 through 6 are 153E9A DNA ClaI digests hybridized with the H7-1, HE340-9, and Kpnl.8 probes respectively. The absence of common fragments demonstrates that these sequences are not closely linked.

These results show that the members of each of the families studied are not significantly intermixed with those of any other family. The alphoid sequences, while found to be
linked, are at least 2 Mb apart. No smaller fragments were observed to which both these sequences hybridized. The smallest fragment to which both the H7-1 and the Kpn1.8 sequences hybridized was 0.56 Mb, indicating a significant distance also lies between these two sequences. The majority of fragments hybridized to only one probe and even most of the fragments greater than 1.0 Mb hybridized to only one probe, indicating that the sequences in this study are not significantly intermixed.

**Double Digests of 153E7BX probed with HH550-31**

Double digests were performed on 153E7BX DNA and probed with HH550-31 to further characterize the organization of this sequence within the centromeric region of the chromosome. The information obtained from double digests permits construction of a more detailed restriction map of the region occupied by this sequence. The double digests performed were BamHI/SalI, ClaI/SalI, BamHI/SfiI, and BamHI/ClaI digests. ClaI, SalI, and SfiI single digests of 153E7BX DNA produce only one band each. Table 7 lists the sizes of the restriction fragments produced by each of the digests and Figure 10 diagrams the information obtained from these digests. This data is preliminary, and the larger fragments (1.03 Mb and up) were in a highly compressed region of the gel where size calibrations were difficult. The numbers presented are the best estimates under the conditions of the gel.
From Table 7, it is apparent that the smallest BamHI fragment (0.32 Mb) is completely contained within the large fragment of the ClaI, SalI or SfiI digests. This follows since this fragment does not change in size from the BamHI single digest to all the double digest involving BamHI. The other two BamHI fragments cannot be placed on the map with the resolution of this gel. In addition, the ClaI/SalI double digest fragment was the size of the ClaI single digest fragment, which on the autoradiogram is slightly smaller than the SalI single digest fragment (Table 7). This observation combined with the 153E7BX data presented in Table 3, which was taken from a gel with better resolution in this region, indicates that the ClaI restriction sites are internal to the SalI sites. The SfiI fragment is the largest of the three; therefore, the SalI sites are internal to the SfiI sites (Table 3). Figure 10 diagrams these restriction sites.

Ribosomal Probe

Because the ribosomal sequences are thought to represent only one to two Mb of the DNA on the short arm of the chromosome, we screened for enzymes that produced fragments containing the ribosomal DNA in one, or at the most two, fragments. The $C_{HB}$ probe was hybridized to 153E9A DNA digested with BamHI, ClaI, PvuII, SalI, SfiI, SstII, XhoI, NotI, MluI, BclI, HpaI, KpnI, BssHII, and SacII. Many of these enzymes are known not to cut within the rDNA repeat
(Sylvester et al., 1986). None of the enzymes tested produced large CHB hybridizing fragments. Figure 9 represents a common result with this probe. 153E9A DNA was digested with BclI and HpaI and run on the low TAFE gel program. All the DNA to which the CHB probe hybridized is compacted at the bottom of the gel, indicating that it is below the resolving range of the gel. This means that the fragments produced by this enzyme, to which the CHB probe hybridize, are less that 0.2 Mb in size. Srivastava and Schlessinger (1991) also observed this type of pattern for ribosomal gene probes, leading them to suggest that the ribosomal genes are not contained in one continuous cluster, but in multiple clusters, interspersed with some yet unidentified sequence(s).
Figure 1. Autoradiograms of restriction digests hybridized with the HH550-31 alphoid probe. 153E9A and 153E7BX DNAs were digested with the indicated enzymes, blotted and hybridized with the HH550-31 alphoid probe. Lanes 1, 3, 5, 7 and 8 are 153E9A DNA digests and lanes 2, 4 and 6 are 153E7BX DNA digests. The fragment sizes in Mb are indicated.
Figure 2. Autoradiograms of restriction digests hybridized with the HE340-9 alphoid probe. 153E9A and 153E7BX DNAs were digested with the indicated enzymes, blotted and hybridized with the HE340-9 probe. Lanes 1, 2, 4 and 6 are 153E9A digests and lanes 3 and 5 are 153E7BX digests. The fragment sizes in Mb are indicated.
Figure 3. Autoradiograms of restriction digests hybridized with the β-satellite probe. 153E9A and 153E7BX DNAs were digested with the indicated enzymes, blotted and hybridized with the H7-1 probe. Lanes 1 through 5 and 7 are 153E9A DNA digests, and lanes 6 and 8 are 153E7BX DNA digests. The fragment sizes in Mb are indicated. In lanes 5 through 8, all the fragments present are below the resolving range of the high gel (≤2.2 Mb).
Figure 4. Autoradiograms of restriction digests hybridized with the satellite III probe. 153E9A and 153E7BX DNAs were digested with the indicated enzymes, blotted and hybridized with the KpnI.8 probe. Lanes 1 through 3 and 5 are 153E9A DNA digests, and lanes 4 and 6 are 153E7BX DNA digests. The fragment sizes in Mb are indicated.
Figure 5. Diagram of chromosome 21 fragments in the somatic cell hybrids used in this study. The cell line 153E9A contains the whole human chromosome 21. The 153E7BX cell line contains the long arm and centromere region of human chromosome 21, and the 2Fu'1 cell line contains only the long arm of human chromosome 21.
Figure 6. Demonstration of linkage between the two alphoid family sequences. Lanes 1 and 3 are probed with HH550-31, and lanes 2 and 4 are probed with HE340-9. Lanes 1 and 2 are 153E9A DNA digested with BamHI, and lanes 3 and 4 are 153E9A DNA digested with Clai. The linkage was confirmed, with both enzymes, by the criteria discussed in the text.
Figure 7. Demonstration of linkage between the satellite III and β-satellite sequences. 153E9A DNA was digested with the indicated enzymes. Lanes 1, 3, and 5 are probed with H7-1, and lanes 2, 4, and 6 are probed with Kpn1.8. The fragment sizes in Mb are indicated. The linkage was confirmed, with all enzymes, by the criteria discussed in the text.
Figure 8. Demonstration of non-linkage between the alphoid and non-alphoid sequences. 153E9A DNA was digested with the indicated enzymes. Lanes 1 and 4 are probed with H7-1, lane 2 is probed with HH550-31, lane 5 is probed with HE340-9, and lanes 3 and 6 are probed with Kpn1.8. The fragment sizes in Mb are indicated.
Figure 9. Autoradiograms of restriction digests hybridized with the ribosomal DNA probe. 153E9A and 153E7BX DNAs were digested with the indicated enzymes. Lanes 1 and 3 are 153E9A DNA digests, and lanes 2 and 4 are 153E7BX DNA digests. The fragments present in lanes 1 and 3 are in the unresolvable portion of the gel and are therefore less than 0.2 Mb.
Figure 10. A partial physical map of the short arm/centromere region of chromosome 21. This diagram illustrates the information obtained in this study. It shows the linkage established between the two alphoid families, and the multiple linkage groups established between the satellite III and β-satellite families, with linkage groups proximal and distal to the rDNA. At least one more cluster of either satellite III or β-satellite exists since there are three linkage groups between these two sequences; however, the identity and location of this cluster is unknown. In addition, it demonstrates that the satellite III sequence was observed to span the breakpoint of the 153E7BX cell line. The 153E7BX double digest information for the alphoid probe HH550-31 is represented by the restriction fragment sites illustrated in the centromere region.
**KEY:**
- **Order Unknown**
- **Linkage Established**
- **153E7BX Breakpoint**
- **TEL** Telomere
- **CEN** Centromere
- **F** = SfiI
- **S** = SalI
- **C** = ClaI
- **B** = BamHI

*Not drawn to scale*
Table 1. Characteristics of the PFGE programs used in this study. Three different pulsed field gel electrophoresis programs were utilized in this study. The parameters of each of the programs are summarized.
**Table 2. Description of the probes utilized in this study.** Five different probes were used in this study. The characteristics of each are indicated.

<table>
<thead>
<tr>
<th>Probe</th>
<th>Family</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH550-31</td>
<td>Alphoid (SX subfamily)</td>
<td>3.25 repeats of the SX 170 bp repeat</td>
</tr>
<tr>
<td>HE340-9</td>
<td>Alphoid (dimer subfamily)</td>
<td>one repeat of the 340 bp dimer repeat</td>
</tr>
<tr>
<td>H7-1</td>
<td>β-satellite</td>
<td>68 bp tandem repeat</td>
</tr>
<tr>
<td>Kpn1.8</td>
<td>Satellite III</td>
<td>5 bp tandem repeat</td>
</tr>
<tr>
<td>C_{HB}</td>
<td>Ribosomal</td>
<td>Specific to the spacer region C</td>
</tr>
</tbody>
</table>
Table 3. Sizes of restriction fragments hybridizing to the alphoid family probe HH550-31. 153E9A and 153E7BX DNAs were digested with several enzymes. The enzymes that produced major bands with the probe HH550-31 are presented in the table, and the total for each of the enzymes is indicated. Other enzymes that were tried but did not produce major bands were KpnI, PvuII, SstII, MluI, XhoI, NotI, BsshII, and SacII. Digests of 2Fu'1 DNA showed no hybridization with this probe. * indicates linkage between the HH550-31 and HE340-9 sequences on these fragments.

<table>
<thead>
<tr>
<th></th>
<th>BamHI (Mb)</th>
<th>ClAI (Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E9A (Mb)</td>
<td>E7BX (Mb)</td>
</tr>
<tr>
<td></td>
<td>3.00*</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>1.60</td>
<td>1.52</td>
</tr>
<tr>
<td></td>
<td>1.52</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td>1.08</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>0.63</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>0.29</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td></td>
<td>8.12</td>
<td>3.41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SalI (Mb)</th>
<th>SfiI (Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E9A (Mb)</td>
<td>E7BX (Mb)</td>
</tr>
<tr>
<td></td>
<td>1.60</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>1.47</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>1.31</td>
<td>4.14</td>
</tr>
<tr>
<td></td>
<td>1.08</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>0.70</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td></td>
<td>6.61</td>
<td>1.47</td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E7BX (Mb)</td>
<td>E9A (Mb)</td>
</tr>
<tr>
<td></td>
<td>8.13</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>8.83</td>
<td>1.62</td>
</tr>
</tbody>
</table>
Table 4. Sizes of restriction fragments hybridizing to the alphoid family probe HE340-9. 153E9A DNA was digested with several different enzymes. The enzymes that produced major bands with the HH340-9 probe are presented above, and the totals are indicated. Other enzymes that were tried but did not produce major bands include KpnI, SalI, SstII, MluI, XhoI, and NotI. Digests of 153E7BX and 2Fu1 DNAs showed no hybridization with this probe. * indicates linkage between the HE340-9 and HH550-31 sequences on these fragments. + indicates that these values were taken from Chawla-Gupta (1991).

<table>
<thead>
<tr>
<th>BamHI (Mb)</th>
<th>ClAI (Mb)</th>
<th>PvuII (Mb)</th>
<th>SfiI (Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.00 *</td>
<td>3.30 *</td>
<td>≥1.40</td>
<td>3.26</td>
</tr>
<tr>
<td>0.64</td>
<td>1.05</td>
<td>0.90</td>
<td>2.70</td>
</tr>
<tr>
<td>0.20 +</td>
<td>0.78</td>
<td>0.40</td>
<td>----</td>
</tr>
<tr>
<td>0.11 +</td>
<td>0.30</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>3.95</td>
<td>5.43</td>
<td>2.70</td>
<td>5.96</td>
</tr>
</tbody>
</table>
Table 5. Sizes of restriction fragments hybridizing to the \( \beta \)-satellite family probe H7-1. 153E9A DNA was digested with several different enzymes. The enzymes which produced major bands with the H7-1 probe are presented above, and the totals are indicated. Other enzymes that were tried but did not produce major bands included BamHI, PvuII, KpnI, XhoI, and NotI. Digests of 153E7BX and 2Fu1 DNAs showed no hybridization with this probe. + indicates linkage between the H7-1 and Kpn1.8 sequences on these fragments.
Table 6. Sizes of restriction fragments hybridizing to the satellite III family probe Kpn1.8. 153E9A and 153E7BX DNAs were digested with several enzymes. The enzymes which produced major bands are presented above, and the totals are indicated. Other enzymes that were tried but did not produce major bands include XhoI, NotI, PvuII, BamHI, and KpnI. Digests of 2Fu1 DNA showed no hybridization with this probe. + indicates linkage between the Kpn1.8 and H7-1 sequences on these fragments.
Table 7. Double digests of HH550-31 alphoid sequences in 153E7BX DNA. 153E7BX DNA was double digested with the restriction enzymes indicated above and hybridized with HH550-31. The single enzyme digests served as controls for this experiment. The sizes of fragments produced in these digests are presented above.

<table>
<thead>
<tr>
<th>BamHI(Mb)</th>
<th>BamHI/SalI(Mb)</th>
<th>SalI(Mb)</th>
<th>ClaI/SalI(Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥1.30</td>
<td>~1.29</td>
<td>~1.29</td>
<td>~1.29</td>
</tr>
<tr>
<td>1.03</td>
<td>1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.32</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SfiI(Mb)</th>
<th>BamHI/SfiI(Mb)</th>
<th>ClaI(Mb)</th>
<th>BamHI/ClaI(Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>~1.29</td>
<td>~1.29</td>
<td>~1.29</td>
<td>~1.29</td>
</tr>
<tr>
<td>1.03</td>
<td>1.03</td>
<td></td>
<td>1.03</td>
</tr>
<tr>
<td>0.32</td>
<td>0.32</td>
<td></td>
<td>0.32</td>
</tr>
</tbody>
</table>
Table 8. Total DNA accounted for by the five probes utilized in this study. The total sizes of the restriction fragments detected by each probe are indicated. The range represents the totals obtained using different enzymes. * indicates that linked fragments exist between the two alphoid probes. + indicates that linked fragments exist between the satellite III and β-satellite probes. Linked fragments were only counted once in the total. † indicates that these values were obtained through calculations from Worton et al., 1988.

<table>
<thead>
<tr>
<th>Family (probe)</th>
<th>Totals of Fragment Sizes Containing Each Family</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alphoid (550)</td>
<td>8.1 - 8.8 Mb*</td>
</tr>
<tr>
<td>Alphoid (340)</td>
<td>4.0 - 6.0 Mb*</td>
</tr>
<tr>
<td>β-satellite (H7-1)</td>
<td>3.8 - 7.2 Mb+</td>
</tr>
<tr>
<td>Satellite III (Kpn1.8)</td>
<td>7.0 - 12.2 Mb+</td>
</tr>
<tr>
<td>Ribosomal gene (C_HB)</td>
<td>1.3 - 1.8 Mb†</td>
</tr>
<tr>
<td>TOTAL</td>
<td>21.2 - 31.9 Mb</td>
</tr>
</tbody>
</table>
DISCUSSION

The short arm/centromere regions of all the human acrocentric chromosomes (13, 14, 15, 21 and 22) have a similar structure consisting of three regions: the centromere-proximal p11 region, the rDNA p12 region, and the distal-telomeric p13 region (Evans et al., 1974). The region is mainly heterochromatic with the ribosomal genes being the only actively transcribed genes (Watkins et al., 1987). Some repetitive DNA sequences are frequently shared by more than one acrocentric chromosome. For example, Vissel and Choo (1991) found four alpha satellite subfamilies that were common to chromosomes 13, 14 and 21, with 13 and 21 being more similar. In another study, Choo (1990) showed similar alphoid sequences present on the short arms of the acrocentric chromosomes; however, the gross organization of these sequences varied between the chromosomes.

These previous studies provided some information about the organization of repetitive sequences on the short arm/centromere region of human chromosome 21; however, no detailed physical map had been constructed for this region of
the chromosome. Chawla-Gupta (1991) determined that the HE340-9 and HH550-31 alphoid families were not intermixed with each other and that they were proximal to the rDNA. This information combined with data from Palamidis-Bourtsos (1989) demonstrated that the HH550-31 family existed in multiple clusters with unrelated sequences interspersed. This previous work also showed that the HE340-9 alphoid family existed as a few, or even a single cluster with no other sequences intermixed. The β-satellite sequences and the satellite III sequences have clusters both proximal and distal to the rDNA (Doering et al., unpublished observation; Greig and Willard, 1992). In combination with this earlier information, the data obtained in the present study allowed construction of the first preliminary overall physical map for the short arm/cen/romere region of chromosome 21. The structural organization revealed in the map provides an explanation for the frequent occurrence of translocations and nondisjunctions involving the acrocentric chromosomes.

Construction of the Physical Map

Clusters of the repetitive sequences. The HH550-31 alphoid, β-satellite, and satellite III sequence probes hybridize to multiple fragments when chromosome 21 DNA is digested with rare cutting restriction enzymes, while the HE340-9 alphoid probe hybridizes to few fragments. This is in agreement with previously reported data (Chawla-Gupta, 1991).
that the HE340-9 alphoid sequence family present on chromosome 21 exists as a single or few clusters while the HH550-31 alphoid sequence exists as multiple clusters. The present data indicate that the β-satellite and satellite III sequences also occur as multiple clusters. Studies on chromosomes 1 and 17 have found the alphoid sequences to be interspersed with other sequences (Carine et al., 1989; Willard and Waye, 1987b). However, studies on the Y chromosome and chromosome 3 found the alphoid sequences to exist as long uninterrupted clusters (Tyler-Smith and Brown, 1987; Willard and Waye, 1987b). Thus, it appears that the organization of alphoid repetitive sequences is very complex and may be chromosome-specific as well as different for the various alphoid families.

None of the restriction enzymes tested in this study produced resolvable fragments when hybridized with C_{HB}, the rDNA probe. Most of the enzymes produced fragments too small to be resolved, indicating that the fragments were less than 200 kb long. The rDNA 44 kb repeat is estimated to have 30-40 copies per acrocentric chromosome, thus occupying 1.3 to 1.8 megabases of DNA on each acrocentric (Worton et al., 1988). The small fragments generated by the restriction digests performed would suggest that the repeating units may not be in a contiguous tandem array. If, as is usually the case for tandem arrays, the repeat units are nearly identical in sequence, a restriction enzyme that does not cut within the
repeat unit should leave the array in one or a few fragments. However, if the rDNA is in a non-contiguous array, it would contain unrelated sequence interspersed with the repeat units. Then, unexpected restriction sites between the repeat units are likely to be present, which, as seen in my results, would cause the rDNA to be cut into a number of smaller fragments. Labella and Schlessinger (1989) have previously suggested that the individual rDNA repeat units are separated by non-ribosomal DNA. They base this conclusion on their inability to isolate yeast artificial chromosomes (YACs) containing more than one repeat unit of the rDNA and on the observation that all the clones also contained unrelated DNA adjacent to the rDNA repeat (Labella and Schlessinger, 1989). In addition, PFGE data has demonstrated that some of the rDNA units are organized in clusters of less than six tandem repeats (Srivastava and Schlessinger, 1991). No matter what enzyme they used, no fragments larger than approximately 200 kb were detected (Srivastava and Schlessinger, 1991), as my results also show. Thus, the size of the region of rDNA could be much larger than previous estimates because of the non-rDNA interspersed within the repeat units (Worton et al., 1988; Sylvester et al., 1986). This would explain the inability to obtain large fragments of rDNA since the non-ribosomal DNA would likely contain additional restriction sites, producing fragments too small to be resolved with the PFGE programs used in this study.
There are two possibilities to explain the nature of this extra sequence. The first is that the rDNA repeat is longer than original estimates, containing previously unknown sequences. The second is that the rDNA repeats have other sequences interspersed within them (Srivastava and Schlessinger, 1991). There are precedents for this second possibility. First, there is evidence in flax that the rDNA may be interrupted by non-rDNA sequence (Agarwal, et al., 1992). A second example is the organization of the HH550-31 alphoid sequence observed in this study and previous studies (Chawla-Gupta, 1991). Another example is in the organization of the β-satellite sequences which appear to be interspersed with two different additional repetitive sequences (Meneveri et al., 1993; McCutheon et al., unpublished observation).

**Localization of the sequences using hybrid cell lines.**

Results using the hybrid cell lines containing chromosome 21 fragments show that the HH550-31 family sequences are both proximal and distal to the 153E7BX breakpoint. This can be concluded since the probe for this family hybridizes to both the 153E9A and 153E7BX cell lines but not to the 2Furl cell line, which contains only the long arm of chromosome 21 (see Figure 5 for a diagram of the cell lines). The HE340-9 family only hybridizes to the 153E9A cell line, indicating that it exists only distal to the breakpoint of the 153E7BX cell line. The Kpn1.8 satellite III probe hybridizes to both the 153E9A
and 153E7BX cell lines indicating that this sequence exists both proximal and distal to the 153E7BX breakpoint. Two different enzymes each produce one satellite III containing fragment that is different in size between the 153E9A and the 153E7BX digests. This indicates that this sequence spans the 153E7BX breakpoint. The portion of chromosome 21 present in the 153E7BX cell line is a translocation to a Chinese hamster ovary chromosome (VanKeuren et al., 1986). Thus, the finding of a breakpoint within the satellite III sequence relates to the recent identification of chromosomal breakpoints in Robertsonian translocations. Wolff and Schwartz (1992) determined that the most frequent breakpoint in the translocations in their study was distal to the centromeric alphoid sequences and proximal to the rDNA and the β-satellite sequences. Gravholt et al. (1992) localized the breakpoints in their study specifically to the satellite III sequences. Therefore, this sequence may represent a weak spot in the genome, prone to breakage.

Since the satellite III sequence spans the 153E7BX breakpoint, the HH550-31 alphoid sequence present on 153E7BX must be more centromeric than the satellite III sequence. In addition, this also supports the existence of multiple clusters of the HH550-31 alphoid family since at least two distinct clusters must exist, one proximal to the 153E7BX breakpoint and one distal to the 153E7BX breakpoint. The β-satellite probe, H7-1, only hybridizes to the 153E9A cell
line, indicating that all of its clusters are present distal to the 153E7BX breakpoint.

**Linkage established between families.** The linkage data provided detailed information on the organization of these sequences in this region of the chromosome. The HH550-31 and HE340-9 alphoid families are linked on a 3.0 Mb BamHI fragment and on a 3.3 Mb ClaI fragment. Since no smaller linked fragments (with the exception of an unconfirmed 2.7 Mb KpnI fragment) were observed, it appears that these two families are separated by at least 2 Mb which allows for a block of alphoid sequence at either end of the fragment. Neither of these two alphoid sequences demonstrated linkage with the other sequences used in this study. When sequences are not found to be linked, estimates on the distance between the sequences are not possible since the sequences could be adjacent but have multiple restriction sites between them.

The β-satellite and the satellite III sequence families were also found to be linked. Two different enzymes produced three fragments each, to which both family’s probes hybridized. The smallest fragment to which both sequences hybridized is 0.60 Mb; therefore, in one of the linkage groups, there is less than 0.60 Mb separating the two sequences. This suggests that multiple linkage groups exist for these two sequences, which is consistent with previous data showing these two sequences in clusters both proximal and
distal to the ribosomal genes (Doering et al., unpublished observation; Greig and Willard, 1992). Thus, linkage groups between these two families likely exist both proximal and distal to the ribosomal genes.

**Description of the Physical Map**

Based on the linkage, localization and cluster data, a partial physical map was constructed for the short arm/centromere region of chromosome 21 (Figure 10). The foundation for the map is based on the observation that the satellite III sequence (Kpn1.8) spans the 153E7BX cell line breakpoint. Since the HH550-31 alphoid sequence does not span the breakpoint, the HH550-31 cluster present in the 153E7BX cell line must be more centromeric than the satellite III sequence. Preliminary work by Doering et al. indicates that the L1.26 alphoid family, which was localized to the long arm side of the centromeric region (Devilee et al., 1986), is in fact HH550-31 (unpublished observation). This observation links the map presented in this study to the long arm of the chromosome. The double digest data for the HH550-31 alphoid sequence on the 153E7BX cell line provided a more detailed map in this region. Since the majority of enzymes tested produced only one fragment with the 153E7BX cell line, it suggests that there is only one cluster of the HH550-31 alphoid sequence in this region. The smallest single fragment produced is the 1.45 Mb ClaI fragment. The 0.32 Mb BamHI fragment indicated
on the map is completely contained within this fragment as determined from the double digest data. The SalI and SfiI restriction sites are just outside the ClaI restriction sites. This suggests a clustering of restriction sites adjacent to a block of repetitive sequences, which implies regions of average sequence composition (non-repetitive sequences) are next to the ends of tandem clusters. This underscores the inconclusiveness of non-linkage data. Two sequences may actually be very close but have a cluster of restriction sites between them.

Since the β-satellite sequence and the satellite III sequences are linked on three separate fragments, this would place one cluster of the β-satellite sequence next to the cluster of the satellite III sequence that spans the breakpoint of the 153E7BX cell line. The HH550-31 and HE340-9 alphoid sequences have previously been determined to exist only proximal to the rDNA (Palamidis-Bourtsos, 1989); therefore, the linked clusters of these two sequences are distal to the β-satellite sequence and proximal to the rDNA, and are separated by approximately two Mb. However, the relative order of the two is not yet known. Since the satellite III and β-satellite sequences are known to exist both proximal and distal to the rDNA (Doering et al., unpublished observation; Greig and Willard, 1992), a linkage group in the region distal to the rDNA is also indicated on the map, but the relative order again is not yet known. In
addition to the two linkage groups indicated for the satellite III and β-satellite sequences, a third group exists. Therefore, an additional cluster of the satellite III or the β-satellite sequence could exist adjacent to one of the indicated linkage groups of these sequences; however, the location of this cluster is not known. A second possibility for the third linkage group is that a separate β-satellite/satellite III linkage group exists that is not immediately adjacent to one of the two linkage groups presented in the map.

Amount of the short arm/centromere DNA accounted for by these sequences. Table 8 presents the total DNA detected by each of the sequence probes utilized in this study as well as the combined total for the probes. To estimate how much of the chromosome 21 short arm/centromere DNA is accounted for by the hybridization of these four probes, the restriction fragments produced by each enzyme with each probe were totaled. The totals vary considerably depending on the enzyme (Tables 3-6). A range was selected from the available enzymes. The lower totals could result from restriction sites being within or between clusters, cutting the sequences into fragments small enough to run off the gel. This could also explain some of the variability in totals between the different enzymes. Totals substantially lower than the other digest totals were not included in the range for that probe.
Some of the very high totals could result from incomplete digestion. The total DNA accounted for by the ribosomal genes is based on calculations by Worton et al. (1988). This was necessary due to difficulties in finding enzymes that would cut the rDNA into only a few fragments (discussed above). This calculation represents a minimum amount of DNA occupied by the ribosomal genes. Including this estimate for the rDNA and the fragments detected by the other probes used in this study, the total DNA accounted for is 21.2 to 31.9 Mb, counting linked fragments only once (Table 8). Again, this range is large due to the variations between enzymes.

The majority of restriction enzymes utilized in this study (excluding BamHI, SfiI and probably also ClaI) are potentially methylation sensitive due to CG sequences in their recognition site. Methylation is especially a concern in heterochromatic regions such as the short arm/centromere region of chromosome 21 (Keshet et al., 1986). This property may inflate the totals for these enzymes due to partial digestions. To allow for such situations, the ranges for the H7-1 and Kpn1.8 sequences (which had the widest ranges) could be based on only the SalI and SstII digests, since the totals of these digests are more consistent with each other. This would bring the combined total range down to 18.1 to 27.2 Mb. Future experiments will be aimed at determining the completeness of the methylation sensitive enzyme digests.
Korenberg and Engels (1978) estimated chromosome 21 to be approximately 60 Mb by a DNA labeling method that estimated its size as a percentage of the total genome. This method, which involves counting grains under a microscope, is not as accurate as physical mapping techniques in which actual fragments of DNA are sized by PFGE and added together. The labeling method most likely is especially inaccurate in heterochromatic regions, such as acrocentric short arms, where the DNA is highly compacted. From Korenberg and Engel's data (1978), Gardiner et al. (1988) estimated the short arm to be approximately 20 Mb. However, the data presented in this thesis indicate that estimate was low, even if the fragments observed in my study account for nearly all of the short arm/centromere DNA. Original estimates for the long arm of the chromosome, around 40 Mb (Korenberg and Engels, 1978; Gardiner et al., 1988), have also been found to be too low. The current estimates, ranging from 45 to 50 Mb, were obtained using more accurate methods such as radiation hybrid, pulsed field, and YAC clone mapping techniques (Chumakov et al., 1992; Gardiner, 1990; Ichikawa et al., 1992).

Implications of this Organization Pattern

The centromere of chromosome 21. Although repetitive sequences seem to be organized in group specific and in chromosome specific patterns, the centromeres appear to have a basic organization common to all human chromosomes.
Manuelidis (1978) observed alphoid DNA sequences at the primary constriction of all human centromeres. Other studies have shown that satellite III DNA sequences are in the centromeric regions, near the alphoid sequences (Cooper et al., 1992 and 1993; Grady et al., 1992; Gravholt et al., 1992; Jackson et al., 1992). Cooper et al. (1992 and 1993) found that the Y chromosome centromere has a block of alphoid DNA and an immediately adjacent block of satellite III DNA in the centromeric region. On chromosomes 13, 14 and 21, Vissel et al. (1992) found an alphoid subfamily to be linked to satellite III DNA sequences. The findings in this study are consistent with these earlier studies. Both the HH550-31 alphoid and satellite III sequence clusters are found in the centromeric region, with the alphoid sequences more centromeric than the satellite III sequences. While these two clusters were not found to be directly linked, it is possible that the alphoid subfamily of Vissel et al. (1992) exists between the HH550-31 alphoid DNA and the satellite III DNA clusters found proximal to the 153E7BX breakpoint (Figure 10). These studies all suggest that there are "basic" features common to the centromeres of all the chromosomes. Thus, these conserved features may play a role in centromere structure and/or function.

Grady et al. (1992) cloned a satellite III sequence that is present at the centromeres of all human chromosomes, and proposed that it exists either interspersed with other
sequences or as distinct uninterrupted clusters. The data in the present study support the second alternative, since the satellite III sequences appear as distinct clusters not frequently intermixed with the other sequences found on the short arm/centromere region of chromosome 21.

Intrinsic properties of the alphoid sequences have led to the hypothesis that they play a role in centromere structure and/or function. These features include their centromeric location and sequence similarity to known functional centromeric sequences in yeast (Fitzgerald-Hayes et al., 1982). It has also been established that they bind the human centromere antigen, CENP-B (Masumoto et al., 1989). Functional roles could include spindle anchors and organizing/sorting sequences in the homologue pairing process or during the formation of the nucleolus. A possible structural role for these sequences is stabilizing the tertiary structure of the chromosome.

The short arm of chromosome 21. The acrocentric chromosomes have been noted to be the most polymorphic in length of the human chromosomes, with the majority of this resulting from variances in the lengths of the short arms (Trask et al., 1989). Chromosome 21 has been observed to be particularly variable in length, with two normal chromosomes varying up to 45%, a range representing 19 Mb (Trask et al., 1989). Since it is not known what the "average" sized
chromosome 21 is, we do not know if the chromosome 21 used in this study, 153E9A, is average. It is not known what type of effect this variance has on homologue pairing and/or separation. This variation has been proposed to result from unequal crossing over between homologues and non-homologues within the various families of repetitive DNA, specifically the tandemly repetitive sequences and the rDNA common to all the acrocentrics (Kurnit, 1979). The multiple cluster organization observed in this study would facilitate unequal crossovers in accordance with the out-of-register recombination model as discussed below (Smith, 1976). Thus, the variation between homologues presumably reflects differences in the copy number of the repetitive DNA present on the short arms/centromeres of these chromosomes (Jacobs, 1977). The assumption with this theory, at least for the crossovers between the non-homologues, is that the acrocentric chromosomes share certain sequences or similar sequences and/or have similar organization of sequences. Studies by Choo (1990) indicate that this is likely a valid assumption. The presence of the rDNA alone sets up a framework for the similar organization of the acrocentrics.

Possible mechanisms of nondisjunction and translocation. The HH550-31, β-satellite, and satellite III sequences, and also possibly the rDNA, exist as multiple clusters. This multiple clusters organization would increase the opportunity
for misalignment of homologues during synapsis, due to greater possibilities for misalignment, resulting in unequal crossing over. If there is only one cluster of a sequence, then there could be only one alignment that would bring the sequences together for crossover. If there are multiple clusters, then clusters in different regions of homologues could line up in several possible ways, creating a misalignment for the whole chromosome. This misalignment could happen in one of two ways between homologues. The first is for the homologues to be misaligned within the same block of a repeat, resulting in a small duplication in one and a deletion in the other homologue. This type of misalignment could happen with single or multiple cluster organization. The second possibility is that different blocks of the same sequence line up, which would be facilitated by the multiple cluster organization observed in this study, resulting in a major duplication on one and a major deletion on the other homologue. Since each of the acrocentrics is thought to have a different organization of similar repetitive sequences, the opportunity for nonhomologous recombination would be enhanced by the presence of multiple clusters of the sequences. The multiple clusters would allow many possible sites of initiation of non-homologous pairing and thus a non-homologous recombination event by the mechanism proposed in the out-of-register recombination model (Smith, 1976).
The acrocentric chromosomes' short arms are in close proximity during the formation of the nucleolus (Jackson-Cook et al., 1985) and accompanying satellite associations (Hansson, 1979). Such non-homologous interactions specifically allow for increased opportunities for translocations since non-homologous chromosomes are in very close contact. If non-homologues begin to pair at homologous regions, it may be detected by the cell as an "error" and could possibly be resolved as a translocation. Again, the presence of multiple clusters of sequence would facilitate such pairings.

Since increased frequencies of structural variations such as double NORs and increased satellite associations have been noted in Down syndrome patients (Hansson, 1979; Jackson-Cook et al., 1985; Jorgensen et al., 1987), it is probable that these variations contribute to nondisjunction and translocation events. For example, delayed or prolonged synapsis between homologues (or non-homologues), due to a structural variation, could result in nondisjunction. It has been hypothesized that an over- or under-representation of a repetitive sequence on the short arm/centromere could result in an increased frequency for nondisjunction and translocation events (Hansson, 1979; Jackson-Cook et al., 1985). Variations could either destabilize or over-stabilize the meiosis I interactions between the homologues. Destabilization would result in the random sorting of the bivalents at the first
meiotic division, whereas over-stabilization would result in the directed mis-assortment of the bivalents. Translocations could result from "reverse" orientation pairing of homologous sequences on non-homologues (Choo et al., 1988). Such a homologous sequence would have an inverted orientation on one of the chromosomes as compared to the other chromosome involved (Choo et al., 1988). Again, the assumption underlying all these possibilities is that the short arms/centromeres of the acrocentric chromosomes have similar organizations of sequences (Trowell et al., 1993; Choo, 1990).

Noted structural variations on chromosomes 21 in Down syndrome patients also adds to the argument that these sequences are not junk DNA. If variations in their structure result in increased frequencies of nondisjunction and translocations, then they most likely function in some manner in the normal meiotic and mitotic processes. Also, the sequences present on the short arm/centromere appear to have a specific orientation within the nucleolus, which also suggests that the sequences are playing some sort of role in the structure and/or function of the nucleolus (Kaplan et al., 1993).

Recently, Hawley and Arbel (1993) proposed a new model for the sequence of events in meiosis I. They propose that the homology search occurs before the actual pairing of the chromosomes. This homology search occurs through single strand out-stretches of DNA sequence, searching for a sequence
with sufficient homology. Through this model, it would be simple to visualize how these sequences could then play a role in the normal pairing process as well as situations resulting in nondisjunction and translocation events. Translocation events could occur if the initial homology search resulted in the mis-pairing of non-homologues. The resolution of this mis-pairing could be a translocation event since double strand breaks in the chromatids begin to appear in preparation for crossing over (Hawley and Arbel, 1993). It is also possible that if this homology search resulted in misalignment of the homologs, the chromatids could undergo unequal crossover and/or become "entangled" resulting in a nondisjunction event, in accordance with the out-of-register model of recombination. Smith (1976) suggested that a minimal degree of sequence homology was necessary for unequal crossovers to be initiated, but it is not known what minimum length of homology is required to facilitate the initial pairing. However, the length of sequence occupied by a repetitive sequence cluster would most likely be sufficient. The multiple cluster organization observed in this study would greatly increase the number of possible erroneous alignments between homologues as well as the mis-pairing of non-homologues.

Understanding the normal organization of these repetitive sequences is imperative to understanding what role or roles they play in nondisjunction and translocation. The pattern of organization observed in this study will eventually be
compared to copies of chromosome 21 in Down syndrome individuals and their parents to determine if there is a risk factor. Any consistent differences found should give some insight as to the role these sequences are playing. It may eventually be possible to use genetic tests to determine predispositions to nondisjunction events. Further studies on acrocentric chromosomes could also facilitate the identification of diagnostic variations in PCD, AD, and cancers, as well as other aneuploidy conditions such as Patau syndrome (trisomy 13). Clearly, further characterization of the structures of the acrocentric short arm/centromere regions will be of great value in defining the roles of these sequences.
BIBLIOGRAPHY


VITA

The author, Jennifer Ann Doll, was born on February 9, 1968 in Kansas City, Missouri. In the fall of 1986, she entered Rockhurst College of Kansas City where she received a B.S. in Biology in May of 1990. In August of 1990, she received a fellowship in the Department of Biology at Loyola University of Chicago to pursue a Master of Science degree which was completed in January, 1994.

In the fall of 1992, Jennifer began pursuing her Ph.D. in Molecular Genetics at Washington University in St. Louis.
PUBLICATIONS


APPROVAL SHEET

The thesis submitted by Jennifer A. Doll has been read and approved by the following committee:

Dr. Jeffrey L. Doering, Director
Associate Professor, Biology, Loyola

Dr. John Smarrelli
Associate Professor, Biology, Loyola

Dr. Holden Maecker
Assistant Professor, Biology, Loyola

Dr. Michael Cummings
Associate Professor, Biological Sciences,
University of Illinois at Chicago

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Science.

10-29-93

Date

Jeffrey L. Doering
Director's Signature