PHYSICAL AND GENETIC MAPPING OF THE X-LINKED
AGAMMAGLOBULINEMIA LOCUS

By

MARIE-ANNE JULIET O'REILLY

A thesis submitted for the Degree of
Doctor of Philosophy (1991)

Institute of Child Health
University of London
ABSTRACT

X-linked agammaglobulinemia is a primary immunodeficiency which arises as a result of a block in B cell differentiation and the subsequent prevention of antibody production by these cells. The disease locus has been previously mapped to Xq21.3-q22 by linkage studies.

Pulsed field gel electrophoresis and radiation hybrid mapping were used to construct a physical map of Xq22. The map incorporated ten probes arranged into three unlinked submaps, and spanned 6 Mb of DNA. The physical map was consistent with genetic recombination frequency data for this part of the genome. PFGE was also used to screen unrelated XLA patients for deletions.

The probe 212/9 (DXS178), which is closely linked to XLA, was found to be flanked by several restriction sites for infrequently cutting restriction enzymes which impaired attempts to span the distance separating the markers which flank the locus. A high resolution physical map was made around this marker which showed that the restriction sites were clustered. This indicated the presence of CpG islands, which are frequently associated with expressed sequences. A chromosome walk towards the CpG islands was instigated by screening a cosmid library with a subfragment of 212/9. Several clones were isolated and two were restriction mapped using infrequently cutting restriction enzymes.

Comparison of the cosmid restriction maps with the genomic map around DXS178 indicated that the cosmids were a maximum of 20 kb from the nearest CpG island.
CONTENTS

Title Page 1
Abstract 2
Contents 3
List of Figures 8
List of Tables 10
Abbreviations 11
Acknowledgements 13

1. INTRODUCTION 14

1.1 Introduction to the X-Linked Immunodeficiencies 14

1.2 Methods used to Investigate the X-Linked Immunodeficiencies 14

1.2.1 Establishment of Disease Loci 15

1.2.1.1 Linkage Analysis 15

1.2.1.2 Analysis of Chromosomal Abnormalities 16

1.2.2 Determination of Affected Cell Lineages using X Chromosome Inactivation Studies 17

1.3 X-Linked Immunodeficiency Diseases (XIDs) 18

1.3.1 X-Linked Severe Combined Immunodeficiency (X-SCID) 20

1.3.2 Wiskott-Aldrich Syndrome (WAS) 21

1.3.3 X-Linked Lymphoproliferative Disease (XLP) 21

1.3.4 X-Linked Hyperimmunoglobulinemia M Syndrome (XHM) 22

1.3.5 Miscellaneous 22

1.4 X-Linked Agammaglobulinemia (XLA) 23

1.4.1 Regional Assignment of the XLA Locus: Data Published Prior to and During the Course of this Study 23

1.4.1.1 Linkage Studies 23

1.4.1.2 Genetic Heterogeneity of XLA 24
<table>
<thead>
<tr>
<th>Section</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4.1.3 Analyses of Patients with Gross Chromosomal Abnormalities</td>
<td>25</td>
</tr>
<tr>
<td>1.4.2 X Chromosome Inactivation Studies in XLA Heterozygotes</td>
<td>26</td>
</tr>
<tr>
<td>1.4.3 XLA with Growth Hormone Deficiency (XLA-GHD)</td>
<td>28</td>
</tr>
<tr>
<td>1.4.4 B Cell Ontogeny in XLA Patients</td>
<td>28</td>
</tr>
<tr>
<td>1.4.5 Gene Candidates for XLA</td>
<td>32</td>
</tr>
<tr>
<td>1.5 Comparison of Human and Murine X-Linked Immunodeficiencies</td>
<td>34</td>
</tr>
<tr>
<td>1.5.1 Candidate Murine Homologues for XLA</td>
<td>35</td>
</tr>
<tr>
<td>1.5.2 Comparison of Human and Murine X Chromosomes</td>
<td>36</td>
</tr>
<tr>
<td>1.6 Positional Cloning</td>
<td>38</td>
</tr>
<tr>
<td>1.6.1 Introduction</td>
<td>38</td>
</tr>
<tr>
<td>1.6.2 Pulsed Field Gel Electrophoresis</td>
<td>40</td>
</tr>
<tr>
<td>1.6.3 Radiation Hybrid Mapping</td>
<td>44</td>
</tr>
<tr>
<td>1.7 CpG Islands, Genes and Methylation</td>
<td>45</td>
</tr>
<tr>
<td>1.7.1 The Base Composition of Vertebrate Genomes is not Uniform</td>
<td>45</td>
</tr>
<tr>
<td>1.7.2 CpG Islands are Usually Associated with Genes</td>
<td>46</td>
</tr>
<tr>
<td>1.7.3 Which Genes have CpG Islands?</td>
<td>46</td>
</tr>
<tr>
<td>1.7.4 What is the Function of CpG Islands?</td>
<td>47</td>
</tr>
<tr>
<td>1.7.5 CpG Islands on the X Chromosome</td>
<td>48</td>
</tr>
<tr>
<td>1.7.6 CpG Islands Contain Sites for Infrequently Cutting Restriction Enzymes</td>
<td>50</td>
</tr>
<tr>
<td>1.8 Aims of this Study</td>
<td>51</td>
</tr>
<tr>
<td>1.8.1 Rational for the Choice of Probes used in this Study</td>
<td>52A</td>
</tr>
<tr>
<td>2. MATERIALS AND METHODS</td>
<td>53</td>
</tr>
<tr>
<td>2.1 Tissue Culture</td>
<td>53</td>
</tr>
<tr>
<td>2.1.1 Lymphoblastoid Cell Lines</td>
<td>53</td>
</tr>
<tr>
<td>2.1.2 Fibroblast Cell Lines</td>
<td>53</td>
</tr>
<tr>
<td>2.1.3 X Chromosome Radiation Hybrid Lines</td>
<td>53</td>
</tr>
<tr>
<td>Section</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>2.1.4</td>
<td>Preparation of Phytohemagglutin (PHA)-Stimulated T Cell Blasts</td>
</tr>
<tr>
<td>2.2</td>
<td>Preparation of DNA in Agarose Blocks for PFGE</td>
</tr>
<tr>
<td>2.2.1</td>
<td>From Blood</td>
</tr>
<tr>
<td>2.2.2</td>
<td>From Sperm</td>
</tr>
<tr>
<td>2.2.3</td>
<td>From Cell-Lines</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Storage of DNA Blocks</td>
</tr>
<tr>
<td>2.2.5</td>
<td>Assessment of DNA Blocks for Degradation</td>
</tr>
<tr>
<td>2.3</td>
<td>Enzymatic Digestion of DNA in Agarose Blocks</td>
</tr>
<tr>
<td>2.4</td>
<td>Preparation of Size Standards for PFGE</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Preparation of Yeast Chromosomes for PFGE Size Standards</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Long Term Storage of Yeast</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Preparation of Lambda Concatamers</td>
</tr>
<tr>
<td>2.5</td>
<td>Pulsed Field Gel Electrophoresis</td>
</tr>
<tr>
<td>2.6</td>
<td>Southern Blotting onto Nylon Filters</td>
</tr>
<tr>
<td>2.7</td>
<td>Prehybridising and Hybridising the Filters</td>
</tr>
<tr>
<td>2.7.1</td>
<td>Prehybridisation</td>
</tr>
<tr>
<td>2.7.2</td>
<td>Hybridisation</td>
</tr>
<tr>
<td>2.7.3</td>
<td>Removal of Non-Specifically Bound Probe</td>
</tr>
<tr>
<td>2.7.4</td>
<td>Autoradiography</td>
</tr>
<tr>
<td>2.7.5</td>
<td>Stripping Filters</td>
</tr>
<tr>
<td>2.8</td>
<td>Transformation of <em>E. Coli</em></td>
</tr>
<tr>
<td>2.8.1</td>
<td>Preparation of Competent Cells</td>
</tr>
<tr>
<td>2.8.2</td>
<td>Transformation of Competent Cells</td>
</tr>
<tr>
<td>2.8.3</td>
<td>Long Term Storage of Transformed <em>E. Coli</em></td>
</tr>
<tr>
<td>2.9</td>
<td>Preparation of Closed Circular DNA using CsCl Gradients</td>
</tr>
<tr>
<td>2.10</td>
<td>Preparation of DNA from Eukaryotic Cells</td>
</tr>
<tr>
<td>2.11</td>
<td>Spectrophotometric Determination of DNA Concentration</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>2.12</td>
<td>Separation of Plasmid Insert and Vector DNA</td>
</tr>
<tr>
<td>2.13</td>
<td>Glass Powder Isolation of DNA from Agarose Gels</td>
</tr>
<tr>
<td>2.14</td>
<td>Quantitation of DNA using Ethidium Bromide Fluorescence</td>
</tr>
<tr>
<td>2.15.1</td>
<td>$^{32}$P-Labelling of DNA using the Random Priming Method</td>
</tr>
<tr>
<td>2.15.2</td>
<td>Removal of Unincorporated $[^a-32]$dCTP using Sephadex-G50 Spin Column Chromatography</td>
</tr>
<tr>
<td>2.16.1</td>
<td>Preparation of High Concentration Sonicated Human DNA for Preannealing Probes Containing Repetitive Sequences</td>
</tr>
<tr>
<td>2.16.2</td>
<td>Preannealing Probes Containing Repetitive Sequences</td>
</tr>
<tr>
<td>2.17</td>
<td>Isolation of DXS178 Positive Cosmids</td>
</tr>
<tr>
<td>2.17.1</td>
<td>Library Screening</td>
</tr>
<tr>
<td>2.17.2</td>
<td>Primary Screening</td>
</tr>
<tr>
<td>2.17.3</td>
<td>Analysis of DXS178 Positive Cosmids using PFGE</td>
</tr>
<tr>
<td>2.18</td>
<td>Reagents</td>
</tr>
<tr>
<td>3.</td>
<td>RESULTS: MAPPING PROBES BY DELETION AND LINKAGE ANALYSIS</td>
</tr>
<tr>
<td>3.1</td>
<td>Localisation of Probes within Xq21.3-Xq22 by Deletion Mapping</td>
</tr>
<tr>
<td>3.2</td>
<td>Linkage Studies with DXS101</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Analysis of Recombinant Chromosomes using DXS101</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Linkage Analysis of XLA Families using DXS101</td>
</tr>
<tr>
<td>3.3</td>
<td>Discussion</td>
</tr>
<tr>
<td>4.</td>
<td>RESULTS: PHYSICAL MAPPING OF Xq22 BY PFGE AND ANALYSIS OF RADIATION HYBRIDS</td>
</tr>
<tr>
<td>4.1</td>
<td>Physical Mapping of Xq22 by Pulsed Field Gel Electrophoresis</td>
</tr>
</tbody>
</table>
4.1.1 Linkage of Probes DXS24, DXS54, DXS83 and PLP

4.1.2 Linkage of DXS17, DXS94, DXS147 and DXS211

4.2 Orientation of Xq22 Probes using Radiation Hybrids

4.3 Attempts to Physically Link DXS178 and PLP to DXS94

4.4 Examination of the Telomeric Deletion Boundary of MBU by PFGE

4.5 Screening XLA Patients for Large Deletions close to DXS178

4.6 Discussion

5. RESULTS: DETAILED PHYSICAL MAPPING AND COSMID CLONING AROUND THE DXS178 LOCUS

5.1 Detailed Physical Mapping around DXS178 using PFGE

5.2 Isolation and Mapping of DXS178 Positive Cosmids

5.2.1 Isolation of a Single-Copy Fragment of DXS178

5.2.2 Isolation of DXS178 Positive Cosmids with 212A

5.2.3 Restriction Enzyme Mapping of Cosmids ICRFc100D1020 and ICRFc100E03167

5.3 Comparison of Genomic and Cosmid Maps around DXS178

5.4 Discussion

6 DISCUSSION - SCOPE FOR FUTURE WORK

BIBLIOGRAPHY

APPENDICES

A1 Details of Probes used in this Study

A2 Primary Screening Details for IFGT Hybrids

A3 Restriction Map of Lorist 4
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The X-Linked Immunodeficiencies</td>
<td>19</td>
</tr>
<tr>
<td>1.2</td>
<td>Deletions in TCD Patients</td>
<td>27</td>
</tr>
<tr>
<td>1.3</td>
<td>Surface Antigen Expression in B Cell Ontogony</td>
<td>29</td>
</tr>
<tr>
<td>1.4</td>
<td>Comparison of the Human and Murine X Chromosomes</td>
<td>37</td>
</tr>
<tr>
<td>1.5</td>
<td>Diagrammatic Representation of a CHEF Apparatus</td>
<td>43</td>
</tr>
<tr>
<td>1.6</td>
<td>Localisation of Probes Used in this Study</td>
<td>52B</td>
</tr>
<tr>
<td>3.1</td>
<td>Localisation of DNA Markers using Deletion Mapping Panels</td>
<td>77</td>
</tr>
<tr>
<td>3.2</td>
<td>Linkage Analysis of Families with Recombinations in Xq21.3-q22</td>
<td>80</td>
</tr>
<tr>
<td>3.3</td>
<td>Linkage Analysis of XLA Families using DXS101</td>
<td>81</td>
</tr>
<tr>
<td>3.4</td>
<td>Refinement of the loci for DXS83, DXS87, DXS114 and DXS178 based on Deletion Mapping and Previous Assignments</td>
<td>83</td>
</tr>
<tr>
<td>4.1</td>
<td>PLP and DXS54 Detect Similar Sized Sal I Fragments</td>
<td>88</td>
</tr>
<tr>
<td>4.2</td>
<td>DXS24, DXS54 and PLP Detect Similar Sized Mlu I Fragments</td>
<td>90</td>
</tr>
<tr>
<td>4.3</td>
<td>Differential Methylation of DNA in Cell Lines LCL-127 and AnLy Produces Different Sized Mlu I Fragments</td>
<td>91</td>
</tr>
<tr>
<td>4.4</td>
<td>DXS24 and DXS83 Detect Similar Sized BssH II Fragments</td>
<td>93</td>
</tr>
<tr>
<td>4.5</td>
<td>PLP, DXS24 and DXS54 Detect Separate BssH II Fragments</td>
<td>95</td>
</tr>
<tr>
<td>4.6</td>
<td>PLP, DXS24 and DXS54 are Located on Separate Sfi I Fragments</td>
<td>96</td>
</tr>
<tr>
<td>4.7</td>
<td>Summary of Physical Linkage Data for PLP, DXS24, DXS54 and DXS83</td>
<td>97</td>
</tr>
<tr>
<td>4.8</td>
<td>Identification of Similar Sized Mlu I Fragments with DXS17, DXS87, DXS94 and DXS147</td>
<td>98</td>
</tr>
<tr>
<td>4.9</td>
<td>Detection of Mlu I Fragments by DXS94, DXS178 and DXS211</td>
<td>99</td>
</tr>
<tr>
<td>4.10</td>
<td>Detection of BssH II Bands with DXS17, DXS87 and DXS147</td>
<td>100</td>
</tr>
</tbody>
</table>
4.11 Detection of BssH II and BssH II/MluH I Bands with DXS94, DXS147 and DXS211  101
4.12 Detection of Sfi I and Sfi I/BssH II Fragments with DXS17, DXS87, DXS94 and DXS147  103
4.13 DXS94 and DXS211 Detect a Similar Sized Sfi I Partial Digest Product  104
4.14 A Physical Map around the Loci DXS17, DXS94, DXS147 and DXS211  105
4.15 DXS94 and DXS178 are Located on Separate Eag I and Nru I Fragments  110
4.16 DXS178 Detects an Altered Eag I Band in MBU  113
4.17 PFGE Analysis of XLA Patients  116
4.18 A Map of Xq22: a Summary of PFGE and Hybrid Mapping Data  121
4.19 Comparison of Homologous Regions on the Human and Murine X Chromosomes Containing the XLA/xid Genes  124
5.1 PFGE Mapping around the DXS178 Locus  130
5.2 A Physical Map Showing Unmethylated Enzyme Sites around the DXS178 Locus  133
5.3 Digestion of ICRFc100D1020 with Infrequently Cutting Restriction Enzymes  138
5.4 Digestion of ICRFc100E03167 with Infrequently Cutting Restriction Enzymes  139
5.5 Restriction Map of ICRFc100D1020  143
5.6 Restriction Map of ICRFc100E03167  144
5.7 Detection of Genomic Kpn I Restriction Fragments with DXS178  145
6.1 Summary: A Physical and Genetic Map of Xq21.3-Xq22  150
<table>
<thead>
<tr>
<th>Table Number</th>
<th>Description</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Frequencies of Restriction Sites for Enzymes in Bulk DNA and in CpG Islands</td>
<td>52</td>
</tr>
<tr>
<td>3.1</td>
<td>Refinement of the Loci for DXS83, DXS87, DXS114 and DXS178 based on Deletion Mapping and Previous Assignments</td>
<td>78</td>
</tr>
<tr>
<td>4.1</td>
<td>Fragments Detected in Female Leukocyte DNA using PFGE</td>
<td>86</td>
</tr>
<tr>
<td>4.2</td>
<td>Fragment Sizes Obtained with <em>Mlu I</em> in Different DNA Samples with PLP, DXS24, DXS54 and DXS83</td>
<td>92</td>
</tr>
<tr>
<td>4.3</td>
<td>Orientation of Xq22 Probes using IFGT Hybrids</td>
<td>107</td>
</tr>
<tr>
<td>4.4</td>
<td>Comparison of Fragments Sizes Obtained with DXS178 by PFGE in MBU and Control DNA</td>
<td>114</td>
</tr>
<tr>
<td>5.1</td>
<td>Restriction Fragments Obtained from Genomic DNA with DXS178</td>
<td>132</td>
</tr>
<tr>
<td>5.2</td>
<td>Restriction Mapping and Hybridisation Data for ICRFcl00D1020</td>
<td>140</td>
</tr>
<tr>
<td>5.3</td>
<td>Restriction Mapping and Hybridisation Data for ICRFcl00E03167</td>
<td>141</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Meaning</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>ADRB</td>
<td>autosomal dominant retinoblastoma</td>
<td></td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
<td></td>
</tr>
<tr>
<td>CF</td>
<td>cystic fibrosis</td>
<td></td>
</tr>
<tr>
<td>CHEF</td>
<td>clamped homogeneous electric field</td>
<td></td>
</tr>
<tr>
<td>CLP</td>
<td>cleft lip and palate</td>
<td></td>
</tr>
<tr>
<td>cM</td>
<td>centiMorgan</td>
<td></td>
</tr>
<tr>
<td>CMGT</td>
<td>chromosome-mediated gene transfer</td>
<td></td>
</tr>
<tr>
<td>DMD/Dmd</td>
<td>Duchenne muscular dystrophy</td>
<td></td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
<td></td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein-Barr virus</td>
<td></td>
</tr>
<tr>
<td>EDA</td>
<td>ectodermal dysplasia, anhydrotic</td>
<td></td>
</tr>
<tr>
<td>EDTA</td>
<td>(Na)₂ ethylenediaminetetraacetic acid</td>
<td></td>
</tr>
<tr>
<td>F9/Cf-9</td>
<td>coagulating factor IX; hemophilia B</td>
<td></td>
</tr>
<tr>
<td>F8/Cf8</td>
<td>coagulating factor VIII; hemophilia A</td>
<td></td>
</tr>
<tr>
<td>FCS</td>
<td>fetal calf serum</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Giemsa</td>
<td></td>
</tr>
<tr>
<td>G6PD/G6pd</td>
<td>glucose-6-phosphate dehydrogenase</td>
<td></td>
</tr>
<tr>
<td>GLA/Ags</td>
<td>alpha galactosidase</td>
<td></td>
</tr>
<tr>
<td>GH</td>
<td>growth hormone</td>
<td></td>
</tr>
<tr>
<td>HAT</td>
<td>hypoxanthine-aminopterin-thymidine</td>
<td></td>
</tr>
<tr>
<td>HPRT/Hprt</td>
<td>hypoxanthine guanine phosphoribosyl transferase</td>
<td></td>
</tr>
<tr>
<td>HYP/Hyp</td>
<td>hypophosphatemia</td>
<td></td>
</tr>
<tr>
<td>IFGT</td>
<td>irradiation and fusion gene transfer</td>
<td></td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
<td></td>
</tr>
<tr>
<td>IL2</td>
<td>interleukin 2</td>
<td></td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pairs</td>
<td></td>
</tr>
<tr>
<td>kD</td>
<td>kiloDalton</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>LM</td>
<td>zone of limiting mobility</td>
<td></td>
</tr>
<tr>
<td>Mb</td>
<td>megabase pairs</td>
<td></td>
</tr>
<tr>
<td>OTC/Otc</td>
<td>ornithine transcarbamylase</td>
<td></td>
</tr>
<tr>
<td>PERT</td>
<td>phenol enhanced reassociation technique</td>
<td></td>
</tr>
<tr>
<td>PFGE</td>
<td>pulsed field gel electrophoresis</td>
<td></td>
</tr>
<tr>
<td>PGK-1/Pgk-1</td>
<td>phosphoglycerate kinase 1</td>
<td></td>
</tr>
<tr>
<td>PLP/Plp</td>
<td>proteolipid protein</td>
<td></td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethylsulphonylfluoride</td>
<td></td>
</tr>
<tr>
<td>RFLP</td>
<td>restriction fragment length polymorphism</td>
<td></td>
</tr>
<tr>
<td>SAR</td>
<td>scaffold associated region</td>
<td></td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
<td></td>
</tr>
<tr>
<td>STS/Sts</td>
<td>steroid sulphatase</td>
<td></td>
</tr>
<tr>
<td>SYN1/Syn1</td>
<td>synapsin 1</td>
<td></td>
</tr>
<tr>
<td>Ta</td>
<td>tabby</td>
<td></td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-borate EDTA</td>
<td></td>
</tr>
<tr>
<td>TCD</td>
<td>tapetochoroidal dystrophy</td>
<td></td>
</tr>
<tr>
<td>TIMP/Timp</td>
<td>tissue inhibitor of metalloprotease</td>
<td></td>
</tr>
<tr>
<td>WAS</td>
<td>Wiskott-Aldrich syndrome</td>
<td></td>
</tr>
<tr>
<td>X-CGD</td>
<td>X-linked chronic granulomatous disease</td>
<td></td>
</tr>
<tr>
<td>X-PD</td>
<td>X-linked properdin deficiency</td>
<td></td>
</tr>
<tr>
<td>X-SCID</td>
<td>X-linked severe combined immunodeficiency</td>
<td></td>
</tr>
<tr>
<td>XHM</td>
<td>X-linked hyperimmunoglobulinemia M syndrome</td>
<td></td>
</tr>
<tr>
<td>XID</td>
<td>X chromosome-linked immunodeficiency</td>
<td></td>
</tr>
<tr>
<td>XLA</td>
<td>X-linked agammaglobulinemia</td>
<td></td>
</tr>
<tr>
<td>XLP</td>
<td>X-linked lymphoproliferative disease</td>
<td></td>
</tr>
<tr>
<td>YAC</td>
<td>yeast artificial chromosome</td>
<td></td>
</tr>
<tr>
<td>ZFX/Zfx</td>
<td>X-linked zinc finger protein</td>
<td></td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I should firstly like to thank my supervisor, Dr. Christine Kinnon for all her help and support over the last three years. I should also like to thank Prof. Roland Levinsky for showing such enthusiasm and support for this project and Dr. Sue Malcolm for helpful contributions.

I should also like to express my gratitude to the Lee Smith Research Foundation for providing my salary for the last two years.

I am grateful to Drs Jean Louis Mandel (Strasbourg), Frans Cremers (Nijmegen), Eddie Bakker (Leiden), Balazs (Lifecodes Corporation, New York) and Karl Tryggvasson (Oulu) for generously providing probes.

I should also like to thank Dr. Dean Nizetic (ICRF, London) for the use of the ICRFc100 cosmid library, and Dr. Peter Goodfellow (ICRF, London) for providing X chromosome radiation hybrids.

The screening of the XLA patients was a collaborative project, and all DNA samples were prepared by Mr. Michel de Weers (Leiden, The Netherlands).

The radiation hybrid mapping was a research idea which I initiated and pursued, however Ms. Lesley Alterman made a major contribution to this area of work and carried out the majority of the Southern blotting. I should therefore like to thank her and also Ms. Karen Curtiss who cultured the radiation hybrid cell lines.

A big thank you also to Ms. Gill Whyman and Ms. Jantien Zjilstra for their contributions over the last three years and to Dr. Stephan Strobel for the use of his Macintosh Computer.

Lastly but by no means least, I should like to thank my family, Jamie and many other friends and colleagues for their invaluable support and encouragement throughout the gestation period of this thesis.
1.1 Introduction to the X-Linked Immunodeficiencies

The components of the immune system include a wide range of tissues and cells and a large number of effector molecules. Defects in the functioning of any of these components can seriously impair the ability of the immune system to function, which is reflected in the fact that over forty different inheritable immunodeficiency diseases have been described in the human in the last four decades (reviewed in Rosen et al., 1984; WHO Scientific Group Report, 1989). The majority of these have autosomal recessive modes of inheritance and therefore the occurrence of symptomatic individuals is restricted to homozygous individuals. In contrast, a few disorders exhibit X-linked recessive modes of inheritance, with female individuals remaining asymptomatic but capable of transmitting the disorder to hemizygous male offspring who manifest the disease. This results in the reappearance of the disease through several family generations. As a result, the few disorders that are X-linked account for a sizeable proportion of all detected pediatric immunodeficiencies and over seventy percent of children with inherited immunodeficiencies are male.

1.2 Methods Used to Investigate the X-linked Immunodeficiencies

The methods used to investigate the X-linked immunodeficiencies are similar to those used to investigate other inheritable disorders. The approaches generally fall into two categories. The first category is that of positional cloning, by which the chromosomal location of a disorder is established and the responsible gene is then cloned on the basis of its chromosomal assignment. The second approach is concerned with deducing the normal function of the responsible gene by comparing the physiology of normal and affected cell populations and (in the case of X-linked disorders) by using X chromosome inactivation studies.
1.2.1 Establishment of Disease Loci

The methods used to establish disease loci include genetic linkage studies and where possible, the analysis of chromosomal abnormalities which occasionally feature in genetic disorders.

1.2.1.1 Linkage Analysis

Genetic linkage studies form part of the primary research carried out on inherited diseases once the criteria for the disease phenotype have been laid down and they have important implications for genetic counselling. The object of such studies is to establish genetic linkage between disease and marker loci by examining meiotic events for crossovers. This is achieved by examining the alleles which have been inherited by offspring of females who are doubly heterozygous for the disease and marker loci.

A prerequisite for a modern genetic linkage study is the prior establishment of a panel of polymorphic DNA probes representing the appropriate chromosome(s). If the disorder exhibits a clear pattern of sex-linked inheritance, probes derived from the X chromosome may suffice; otherwise probes which represent all 22 autosomes may have to be used.

There are currently approximately 230 polymorphic probes listed as mapping to the X-chromosome (Mandel et al., 1989; Williamson et al., 1990). Many of these have been localised using somatic cell hybrid panels (eg. Wieacker et al., 1984; Oberle et al., 1986) and by direct in situ hybridisation to metaphase chromosome spreads (eg. Riddell et al., 1986). A number of linkage studies on the X chromosome have facilitated the construction of linkage maps which represent the entire X chromosome (Drayna et al, 1984; Drayna and White, 1985). In addition, more precise linkage maps of particular regions of the X chromosome have been constructed (Bakker et al., 1986; Bertelson et al., 1986; Oberle et al., 1987; Veenema et al., 1987; Arveiler et al., 1987).

The degree of success achieved with linkage studies depends on the number of polymorphic probes used and on the frequency of the alleles that they
recognise in the population. The majority of polymorphic DNA probes recognise two alleles which are identified by digesting DNA with a specific restriction enzyme whose restriction site has been created/destroyed in a proportion of the population. The likelihood of a female carrier being heterozygous with a particular probe will therefore increase as the allele frequencies tend towards 50%.

Obviously, the number of informative meioses studied also affects the strength of the linkage data. The likelihood that two loci are linked is expressed as a lod score (see Ott, 1986), where lod (Z) equals the log₁₀ of the ratio of the odds that two loci are linked to the odds that they are unlinked. Genetic distance (Θ) is expressed as the percentage recombination between two loci. A lod score of 3 or greater is taken to be positive evidence of linkage.

The map units of genetic distance are centiMorgans (cM), where 1 cM is approximately equal to a 1% recombination frequency between two loci. However, the relationship between cM and Θ changes as Θ increases because of the likelihood of double recombination events between two loci. This relationship is described by a mapping function. The recombination fraction for two unlinked probes is Θ = 0.5.

The length of the human genome is 33 Morgans (M). Since the size of the haploid genome is approximately 3×10⁹ bp, a correlation can be drawn between physical and genetic distance, so that 1 cM is approximately equivalent to 1 Mb of DNA. This relationship cannot be used to accurately convert recombination fraction into physical distance since recombination varies over the chromosomes.

It is also possible to calculate the relative orders and genetic distances between groups of probes assigned over several cM using multipoint linkage analysis, with the aid of computer programs such as LINKAGE (Lathrop et al., 1984).
1.2.1.2 Analysis of Chromosomal Abnormalities

A small proportion of genetic disorders are due to large chromosomal deletions which are cytogenetically detectable. There are two regions of the X chromosome which are associated with large male-viable deletions. These are the large, dark staining Giemsa (G) bands, Xp21 (Francke et al., 1985; Old et al., 1985) and Xq21 (Hodgson et al., 1987; Schwartz et al., 1988; Cremers et al., 1988). This is consistent with the apparent enrichment for expressed sequences of Giemsa lightstaining bands (or reverse bands) compared to the "genetic inertia" of G bands (Korenburg et al., 1978).

1.2.2 Determination of Affected Cell Lineages using X Chromosome Inactivation Studies

In order to determine which cell lineages are affected in the various disorders, X chromosome inactivation studies have been used. These are based on the observation that early in embryogenesis, the random inactivation of one of the two chromosomes in female somatic cells occurs (Lyon, 1972). This inactivation is transmitted in stable fashion to all cell progeny and only genes from the active X chromosome can then be expressed. In female carriers of X-linked immunodeficiencies, cells from lineages in which the product of the defective gene plays a primary developmental role will be arrested during differentiation. As a result, all cells from the affected lineages which mature normally will express the non-defective gene and therefore will show unilateral inactivation of the affected chromosome.

X chromosome inactivation patterns can be determined in three ways. The original method utilised the isoforms of the enzyme glucose-6-phosphate dehydrogenase (G6PD) (Nyhan, 1970; Prchal et al., 1980). This method has limited applications, however, because the G6PD protein polymorphism occurs in less than 1% of the Caucasian population.

The second method distinguishes active and inactive X chromosomes on the
basis that the latter show hypermethylation of particular 5' -CpG- 3' rich areas of DNA which are hypomethylated on the active X chromosome (Wolf et al., 1984a; 1984b; Battistuzzi et al., 1985; Keith et al., 1986). The restriction enzyme \textit{Msp} I, which recognises the sequence CCGG and its isoschizomer \textit{Hpa} II, which is methylation-sensitive, can be used to digest DNA from particular cell lineages. This DNA can then be size separated and hybridised with probes spanning particular CpG-rich regions. If the X inactivation pattern is random, the CpG rich regions of approximately half the DNA will be digested by \textit{Hpa} II whereas \textit{Msp} I will digest both chromosomes. The resulting band patterns will reveal the randomness of X chromosome inactivation.

A third method of determining X inactivation status is by fusing human cells with the Chinese hamster fibroblast line RJK88 or the murine myeloma X63-Ag8.653 which are both deficient for the X-linked gene hypoxanthine guanine phosphoribosyl transferase (HPRT) (Puck et al., 1987; Hendriks et al., 1989). After fusion, the cells are grown in HAT (hypoxanthine-aminopterin-thymidine) which selects for cells with an active HPRT gene. Subsequent analysis of the hybrids for any informative X-linked RFLP will identify the X chromosomes in the surviving hybrids. At least ten hybrids must be produced to provide a representative sample.

1.3 X-linked Immunodeficiency Diseases (XIDs)

So far, seven X-linked immunodeficiency diseases (XIDs) have been defined (see Figure 1.1). Two affect non-specific immunity: these are X-linked chronic granulomatous disease (X-CGD), which affects phagocyte function and X-linked properdin deficiency (X-PD) which affects the alternative complement activation pathway. The genes for both of these immunodeficiencies have been cloned. The X-CGD gene codes for the \(\beta\)-chain of cytochrome \(b_{245}\) (Teahan et al., 1987; Dinauer et al., 1987) and is situated on the short arm of the X chromosome at Xp21.1. The X-CGD gene was cloned on the basis of its
Figure 1.1
The X-Linked Immunodeficiencies

Xp
22.3
22.2
21.3
21.2
21.1
11.4
11.3
11.23
11.22
11.21
11.2
12
13
X-SCID

Xq
21.1
21.2
21.3
22.1
22.2
22.3
23
24
25
26
27
28
XLA
XHM
XLP

X-PD | X-CGD
WAS
chromosomal location which had been determined using linkage and deletion analysis (Royer-Pokora et al., 1986).

X-PD has been localised to the short arm of the X chromosome in the region Xp11.23-p21.1. Properdin is a serum protein which promotes activation of the alternative complement pathway. The gene for properdin has recently been cloned (Nolan et al., 1991) and it remains to be seen whether mutations within the gene will be found in X-PD patients.

Five of the defined X-linked immunodeficiencies affect the lymphoid cell lineages. These are X-linked agammaglobulinemia (XLA), X-linked severe combined immunodeficiency (X-SCID), Wiskott-Aldrich Syndrome (WAS), X-linked lymphoproliferative disease (XLP), and X-linked hyper-immunoglobulinemia M syndrome (XHM). As yet, no candidate genes or gene products have been identified for any of these disorders although they have been defined as discrete syndromes using the investigatory methods which have been described above. The XIDs are described in brief below.

1.3.1 X-Linked Severe Combined Immunodeficiency (X-SCID)

Severe combined immunodeficiency (SCID) is a disorder characterised by the absence of T cell-mediated and humoral immunity. SCID is fatal by 1 to 2 years of age if not treated by bone marrow transplantation. Two thirds of SCID cases are caused by a single X-linked gene, which has been mapped by family linkage studies to the long arm of the X chromosome at Xq11-q13 (de Saint-Basile et al., 1987; Goodship et al., 1989a; Puck et al., 1990). Using multipoint analysis a likely probe order has been defined as centromere - DSX159 - (PGK1, DXS347, DXS441, DXS447, X-SCID) - DXS72 - telomere. No recombinations have occured between X-SCID and PGK1, DXS347, DXS441 or DXS447 in over 20 informative meiosis.

Research into the X chromosome inactivation patterns of females who carry X-SCID has demonstrated that in B cells, T cells and granulocytes there is a non-random pattern of X chromosome inactivation (Puck et al., 1987; Conley et
al., 1988; Goodship et al., 1988). This indicates that the gene may be expressed in all these cell types.

1.3.2 Wiskott-Aldrich Syndrome (WAS)

Wiskott-Aldrich Syndrome is characterised by defective cellular and humoral immunity, thrombocytopenia and eczema. Affected males suffer from recurrent, chronic bacterial and viral infections and usually die in the first decade of life. Lymphocytes are present in the circulation of affected males; however they have defective O-glycosylation of the membrane sialoglycoprotein CD43 (Remold-O’Donnell et al., 1984; Greer et al., 1989a), a molecule which has recently been shown to enhance antigen-specific T cell activation (Parks et al., 1991). They also have abnormal microvilli (Kenney et al., 1986). Immunoglobulin levels are abnormal, with high IgA and IgE, low IgM and normal or elevated IgG levels. B cells make no antibody response to polysaccharide antigens.

X chromosome inactivation studies have demonstrated non-random inactivation in platelets, granulocytes, monocytes, and B and T cells of females carriers (Fearon et al., 1988; Greer et al., 1989b). As normal levels of these cell populations are found in males who manifest the disease, the WAS defect must impart a severe but not absolute growth disadvantage to these cells.

The WAS locus has been assigned to the region Xp11.4-p11.21 (Peacock and Simnovitch, 1987; Kwan et al., 1988; de Saint Basile et al., 1989; Greer et al., 1990). No recombinations have been found with the hypervariable probe DXS255 in over 40 informative meioses.

1.3.3 X-Linked Lymphoproliferative Disease (XLP)

Affected males with this disorder are unable to make an antibody response to Epstein-Barr virus (EBV) nuclear-associated antigen. This can result in severe or fatal infectious mononucleosis, hypogammaglobulinemia, and malignant lymphocytic proliferations (Purtilo et al., 1975). The EBV receptor is present
on normal B-lineage cells from a precursor stage through differentiation to antibody-secreting plasma cells. The wide range of developmental stages during which these cells could become infected may account for the large degree of heterogeneity in the phenotype of this disorder. X chromosome inactivation studies in two families have demonstrated random inactivation in B cells and T cells (Neidich et al., 1989). XLP has been assigned to Xq24-q27 by linkage studies (Skare et al., 1989). This data has been substantiated by the identification of an interstitial deletion within Xq25 in an XLP patient (Wyandt et al., 1989).

1.3.4 X-Linked Hyperimmunoglobulinemia M Syndrome (XHM)

Patients with X-linked hyperimmunoglobulinemia M syndrome suffer from frequent, chronic infections. They are also susceptible to autoimmune haemolytic anaemia and hepatomegaly or lymphadenopathy (Rosen and Janeway, 1966). XHM syndrome is characterised by selective agammaglobulinemia of the IgG and IgA antibodies. In contrast, IgM levels are frequently abnormally high and non-functional. B cell numbers appear to be normal as do T cell numbers and function. The mechanism that enables B cells to switch from production of IgM to that of other antibody classes appears to be intact. X chromosome inactivation studies in female carriers have demonstrated random patterns of inactivation in B cell populations producing IgA or IgG antibodies (Hendriks et al., 1990). Linkage has been found to the polymorphic DXS42 locus at Xq24-q27, which is the same region to which XLP has been mapped (Mensink et al., 1987).

1.3.5 Miscellaneous

Recently, an additional locus on the X chromosome has been demonstrated to be of immune significance. An extended family was found to contain several males with symptoms including recurrent sinusitis, otitis media, bronchitis and pneumonia; severe varicella and chronic papillomavirus infections. Investigations
revealed that the abnormalities were mainly confined to T cell number and function and were distinct from the phenotypes of any previously characterised X-linked immunodeficiency (Brooks et al., 1990).

1.4 X-Linked Agammaglobulinemia (XLA)

XLA was the first X-linked immunodeficiency to be described (Bruton, 1952). Affected males have recurrent and protracted bacterial infections, as a result of extremely low serum Ig levels. Peripheral B cells are virtually absent (Conley, 1985).

1.4.1 Regional Assignment of the XLA locus: Data Published Prior to and During the Course of this Study

1.4.1.1 Linkage Studies

The XLA locus has been assigned to Xq21.3-q22 on the basis of its close linkage with probes which map to this region (Kwan et al., 1986; Mensink et al., 1986a; Malcolm et al., 1987; Lau et al., 1988; Guioli et al., 1989, Kwan et al., 1990).

In 1986, Kwan et al. produced linkage data which indicated that the XLA gene was located on the proximal long arm of the X chromosome, close to the loci DXS3 (Z=3.65 at Θ=0.04) and DXS17 (Z=2.17 at Θ=0). In addition, Mensink et al. (1986a) found linkage with DXS3 in a single large pedigree (Z=3.30 at Θ=0.06). Malcolm et al. (1987) found that XLA was closely linked to DXS94 (Z=6.6 at Θ=0: 95% confidence intervals 0 < Θ < 0.102) and DXS17 (Z=4.4 at Θ=0: 95% confidence limits 0 < Θ < 0.157). One recombination was found between DXS3 and XLA (Z=3.63 at Θ=0.05). Until 1987, the relative orders of probes in the region Xq13-q22 had not been determined. Arveiler et al. (1987) constructed a linkage map of proximal Xq using a total of 44 pedigrees, including 17 from the Centre d'Etude du Polymorphisme Humain (CEPH: Dausset et al., 1990). The map indicated that
the most likely order of loci was centromere-DXS3-(DXS94, DXS178)-DXS17- telomere. In two-point linkage tests, the recombination fractions that gave maximum lod scores were DXS3-0.04-DXS94; DXS3-0.03-DXS178; DXS3- 0.10-DXS17; DXS94-0.00-DXS178; DXS94-0.02-DXS17; DXS178-0.02- DXS17.

The probe DXS178, which showed close linkage to DXS94 (Z = 10.31 at θ=0; 95% confidence limits 0 < θ < 0.061) and DXS17 (Z = 11.86 at θ=0.01) in the study by Arveiler et al. (1987), has recently been used in linkage studies in XLA families. Guioli et al. (1989) found no recombinations between XLA and DXS178 (Z = 5.92 at θ=0.00); similarly Kwan et al. (1990) found no recombinations between the two loci in over 15 informative meioses and Malcolm et al. (1989) found no recombinations in a further twelve meioses. The combined two-point lod score from these studies is Z = 14.48 at θ=0 with 95% confidence intervals of 0 < θ < 0.031.

The data from Arveiler et al. (1987), combined with that from the XLA linkage studies, indicate that DXS3 flanks XLA on the centromeric side. The data published by Kwan et al. (1990) has made it possible in addition to determine the relative order of XLA and the two closely linked markers, DXS94 and DXS178. Kwan et al. have used the order of markers from Arveiler et al. (1987) with their linkage data. They have the most likely order of probes, with recombination fractions based on multipoint linkage analyses as centromere-DXS3-6.5cM-(XLA, DXS178)-5cM-DXS94-3.5cM-DXS17-telomere. As no recombinations between DXS178 and the disease have been detected to date, it is not possible to deduce the relative order of XLA and this marker.

1.4.1.2 Genetic Heterogeneity of XLA

There have been reports of XLA families where the linkage data has given negative lod scores between XLA and probes from Xq22, suggesting the possibility of a second location for the mutated gene in these families. In one such pedigree, two sisters, each with affected sons, showed unilateral X chromosome inactivation in the B cell lineages (Ott et al., 1986; Mensink et al., 1986a). However, the chromosome that was inactivated was found to be derived
from their father, not their mother. As their father was healthy and XLA has full penetrance, this indicates that he was an X chromosomal mosaic, i.e. his B cells contained an intact XLA gene whereas his spermatocytes contained an X chromosome with a defective XLA gene. When the linkage analysis was repeated using a model of paternal transmission, the pedigree showed no recombinations between XLA and DNA markers from Xq22 (Hendriks et al., 1989). A second family which gave negative lod scores with Xq22 markers was later diagnosed as having XHM rather than XLA (Lau et al., 1988; Timmers et al., 1991). This data is consistent with a single XLA locus at Xq21.3-Xq22. Germline mosaicism has been recorded previously in other disorders such as Duchenne muscular dystrophy (DMD) and WAS (Darras and Francke, 1987; Arveiler et al., 1990).

Approximately 30% of the XLA patients presenting represent recent mutations. These appear to have arisen during spermatogenesis in the grandfathers in a proportion of cases, for example in 2 out of 13 families reported by Lau et al., 1988.

1.4.1.3 Analyses of Patients with Gross Chromosomal Abnormalities

A number of X-linked male-viable deletions and duplications have been reported to involve the long arm of the X chromosome around the Xq21 region (see Section 1.2.1.2). The disorders reported in these patients have included tapetochoroidal dystrophy (TCD), deafness, cleft lip and palate (CLP) and mental retardation (Cremers et al., 1989). However, none of these patients have been reported to be immunodeficient.

Deletion mapping, using DNA from blood or lymphoblastoid cell-lines originating from these patients, has been used to divide Xq21 into subregions on the basis of the number of copies of probes detected in these patients (Cremers et al., 1987; Cremers et al., 1988; Schwartz et al., 1988; Cremers et al., 1989). Several of these chromosome abnormalities have been shown to extend into
Xq21.3. This region is part of the division to which XLA has been assigned in linkage studies.

In family studies, the probe DXS3 has been shown to be tightly linked to XLA but to lie a few cM centromeric to the XLA locus. No probes have been placed between DXS3 and XLA by multipoint analysis, therefore this probe is the centromeric limit to the XLA locus as determined by linkage analysis.

DXS3 has been deleted in several of the TCD patients including NP (Cremers et al., 1988; Schwartz et al., 1988), RvD and MBU (Cremers et al., 1989) (see Figure 1.2). It follows therefore, from deletion mapping, that the deletion breakpoints in these patients can be considered to be the centromeric boundary to the XLA locus, rather than DXS3.

Deletion mapping has arranged the probes that map distal to DXS3 into two subregions. Two patients, RvD and MBU are deleted for the probes DXS96, DXS118, pF1 and pF8 which lie distal to DXS3. Neither of these patients has XLA, therefore the XLA locus must lie distal to these probes. Patients RvD and MBU have a single copy of several probes present in Xq22 and the distal breakpoints of the deletions in these two patients have not yet been distinguished using the DNA probes shown in Figure 1.2 (Cremers et al., 1988; 1989). Cytogenetic studies have determined that the NP deletion spans Xq21.1-q21.33, with a small portion of Xq21.1 or Xq21.33 still present (Schwartz et al., 1988). The RvD deletion spans Xq21.1-q22.1 (Cremers et al., 1988). The MBU deletion has been described as del(X) Xq21.2-q21.33 (Cremers et al., 1989).

1.4.2 X Chromosome Inactivation Studies in XLA Heterozygotes

X chromosome inactivation studies using G6PD isoenzymes (Conley et al., 1986), DNA methylation studies (Fearon et al., 1987) and RFLP analysis of somatic cell hybrids (Conley and Puck, 1988; Hendriks et al., 1989) have all demonstrated that the peripheral B cells of female carriers have unilateral inactivation of the affected X chromosome. In contrast, T cells, monocytes and granulocytes in these individuals demonstrate random X chromosome
Figure 1.2
Deletions in TCD Patients

<table>
<thead>
<tr>
<th></th>
<th>Normal X</th>
<th>MBU</th>
<th>NP</th>
<th>RvD</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xq</td>
<td>21.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Markers:
- PGK1
- DXS72
- DXS121
- DXYS1
- DXYS12
- DXS3
- DXS96, DXS118
- pF1, pF8
- DXS17
- DXS94
inactivation. These findings indicate firstly that the XLA gene plays a critical role in the differentiation of precursor B cells and secondly that the primary XLA defect does not involve other hematopoietic cell lineages.

1.4.3 XLA with Growth Hormone Deficiency (XLA-GHD)

There have now been three reports of families containing males who manifest X-linked agammaglobulinemia in association with growth hormone (GH) deficiency (Fleisher et al., 1980; Conley et al., 1989; Sitz et al., 1990). Human growth hormone is encoded on an autosome (Owerbach et al., 1980) so the defect may reside in some regulator of GH. Conley et al. (1989) reported that no cytogenetic abnormalities had been found to provide evidence for a contiguous gene deletion syndrome, although RFLP analyses of this family was consistent with the locus being the same as that of XLA and X chromosome inactivation studies in female carriers demonstrated the same patterns of X inactivation as is seen in classic XLA carriers.

1.4.4 B Cell Ontogeny in XLA Patients

XLA is characterised by agammaglobulinemia involving all classes of immunoglobulin. This appears to result directly from the absence of antibody-producing peripheral B cells. The exact stage at which B cell ontogeny is halted in XLA patients appears to vary slightly.

Several discrete stages in B cell development have now been defined with regard to cell surface antigens and immunoglobulin rearrangement events (reviewed in Uckun, 1990) (see Figure 1.3). However, the regulatory agents which control the progression of B cell precursor development are not well understood. The direct precursors of surface IgM positive (+), peripheral B cells are pre-B cells (Figure 1.3). Pre-B cells are characterised by the presence of cytoplasmic $\mu$ heavy (H) chain protein but the absence of surface IgM. These cells constitute approximately 6% of lymphoid cells in normal bone marrow. The bone marrows of XLA patients have been shown to contain comparable
Figure 1.3
Surface Antigen Expression in B Cell Ontogeny*

<table>
<thead>
<tr>
<th></th>
<th>Pro-B</th>
<th>Pre-pre-B</th>
<th>Pre-B</th>
<th>Immature B</th>
<th>Mature B</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHC-II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IgM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IgD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*adapted from Uckun, 1990.
levels of cytoplasmic $\mu^+$ pre-B cells (4%: Pearl et al., 1978). The stage at which B cell maturation is arrested appears to vary slightly between XLA patients as early surface $\mu^+\delta^+$ B cells have been detected in some patients (Schwaber et al., 1978). This may reflect the degree to which the function of the XLA gene product has been abrogated by different mutations.

The efficiency of the B cell population in recognising foreign antigen stems from the considerable diversity that results from the ingenious usage of the many variable (V), diversity (D) and joining (J) gene segments during rearrangement of the Ig loci (reviewed in Alt et al., 1986). The order of gene rearrangement has been demonstrated to be a) D to $J_H$ and then b) $V_H$ to the rearranged D-$J_H$ segments. This is then transcribed together with the constant (C) region gene, spliced and translated, producing $\mu$ H chain protein. B cell lines from XLA patients have been found to express immature forms of IgH chain molecules comprising of LS-D-$J_H$-$C\mu$ where LS is a leader sequence (Schwaber et al., 1986). These immature products are not unique to XLA patients: they have also been found in mouse Abelson virus-transformed pre-B cell lines (Reth and Alt, 1984). Schwaber (1986) proposed that these immature protein products acted as regulatory elements to control the steps in $V_H$-$D$-$J_H$ recombination, and that the XLA defect affected an X-linked product which regulated Ig gene rearrangement, possibly a recombinase, a conformation regulatory product, or a defect in the promoter site to which the immature products should bind. Anker et al. (1989) have put forward several lines of evidence that oppose the recombinase theory, however. Firstly, both T cells and B cells appear to utilise a common recombinase system (Yancopoulos et al., 1986) so T cell mediated responses would be affected in XLA patients if a recombinase was the primary defect; however, T cell responses are normal in XLA patients (Rosen et al., 1984). Secondly, two EBV-transformed cell lines from an XLA patient have produced native-sized $\mu$ chain protein which can be expressed on the cell surface following fusion with Ig-light chain (L)-producing lymphoblastoid cells (Fu et
al., 1980; McCune and Fu, 1981). Thirdly, $V_H DJ_H$ joining has been shown to occur normally in an EBV-transformed B cell line from an XLA patient (Mensink et al., 1986b) and finally, the small numbers of circulating B cells that are occasionally found in XLA patients have phenotypes characteristic of immature B cells (Tedder et al., 1985). If the XLA defect was limited to Ig rearrangement one would expect the B cells to mature normally once they had passed this point, instead of remaining at an immature stage.

Anker et al. (1989) analysed the Ig loci of 18 EBV-transformed B cell lines from XLA patients. The results indicated that Ig rearrangement of both $H$ and $L$ chain loci occurred normally and with diverse usage of the various gene segments. The fact that these cells were able to rearrange their $L$ chain genes was very interesting, considering the fact that XLA is characterised by the near absence of peripheral sIgM+ cells and there is a requirement for L chain (or L chain substitutes) before Ig can be transported to the cell surface. The XLA pre-B cell lines did not show restricted usage of $V_{\kappa}$ segments (Conley et al. 1989), again suggesting that complete and unrestricted rearrangements were able to occur.

The fact that detectable amounts of IgG can be found in the serum of many XLA patients (Conley, 1985) indicates that not only can these cells make complete Ig molecules, but they are also capable of class switching. These findings suggest that the XLA gene defect is unlikely to be a recombinase enzyme, however the possibility that it is another factor which is involved in the regulation of Ig gene rearrangement events can not be ruled out.

An alternative suggestion for the role of the XLA gene product is as a factor in proliferation or clonal expansion of cells of the B cell lineage. Pearl et al. (1978) used thymidine incorporation and autoradiography to measure the proliferation rate of normal and XLA pre-B cells. The percentage of cytoplasmic $\mu+$, surface $\mu-$ negative cells which had incorporated thymidine was less (21-43%) in XLA patients than in controls (mean = 91%). Also, Conley et al. (1989) noted that EBV-transformed cell lines from XLA patients multiplied at
a greatly retarded rate compared to EBV lines from normal individuals.

1.4.5 Gene Candidates for XLA

Any gene which is expressed in B lineage cells specifically around the pre-B cell and perhaps later B cell stages must be considered for candidature for the XLA gene.

As has been discussed, pre-B cells rearrange their immunoglobulin H chain loci using a recombinase enzyme(s). Free H chain proteins are retained in the endoplasmic reticulum. Subsequently, the cells rearrange their L chain V region loci from their \( V_L \) and \( J_L \) components. Fully assembled H and L chains are transported to the cell surface, leading to the generation of surface IgM+ B cells. It has generally been believed that H chains cannot be expressed on the cell surface unless they are associated with one or other of the kappa or lambda L chains. However there is now evidence that "surrogate" L chains may also function in a similar way in pre-B cells. Sakaguchi and Melchers (1986) identified lambda 5, a murine protein with strong homology to the constant (C) region of the lambda L chain which appeared to be selectively expressed in murine pre-B cells. A second murine gene, \( V_{\text{preB}} \) (Kudo and Melchers, 1987) which has strong homology to the V region of the lambda L chain genes, was found to have a similar pattern of expression. The human equivalents to both of these proteins have been isolated (Kudo et al., 1989). A monoclonal antibody against the human \( V_{\text{preB}} \) chain was found to co-precipitate three proteins with sizes corresponding to the \( V_{\text{preB}} \), lambda 5 and \( \mu \) chain proteins, suggesting that \( V_{\text{preB}} \) and lambda 5 may be required for secretion of \( \mu \) chain. Transfection experiments with the murine genes (Tsubata and Reth, 1990; Karasuyama et al., 1990) indicated that expression of these genes was sufficient to permit cell-surface expression or secretion of \( \mu \) chain. The lambda 5 and \( V_{\text{preB}} \) chains are located on the autosomes (Schiff et al., 1990), however it is feasible that they may play a regulatory role in B cell development involving other molecules.
This is supported by evidence that the early expression of membrane $\mu$ chain (possibly in association with these molecules) is required for pre-B cell differentiation (Kitamura et al., 1991). The pattern of expression of lambda 5 and $V_{\text{preB}}$ in B lineage cells from XLA patients has yet to be elucidated.

A number of other proteins have been found to be associated with H chain expression in the mouse. Takemori et al. (1990) studied a transformed immature B cell clone (Ig6.3) that had surface $\mu$ chains but had not rearranged either of its L chain loci. Immunoprecipitation of cell lysates with anti-$\mu$ antiserum revealed chains of 15.5, 27, 40 and 50 kD in addition to a polypeptide of 14 kD which was probably the $V_{\text{preB}}$ chain. A second pre-B cell clone (Ig6.11) which had rearranged its kappa L chain locus was also found to express the 14, 15.5 and 50 kD proteins, in association with $\mu$ chain but not with kappa L chain. Splenic B cells were not found to express any of these polypeptides. In both cell lines, cross-linking experiments with anti-$\mu$ antiserum produced a rise in intracellular calcium concentrations \([Ca^{2+}]\). A significant increase in inositol phosphate metabolism was seen in the more differentiated cell line, consistent with observations seen in mature B cells; however no increase was detected in the Ig6.3 line. This data indicated that the $\mu$ chain complexes were capable of acting as signal transduction molecules.

An additional protein which is expressed in murine pre-B cells and mature B cells is B34. This appears to play an important role in transporting $\mu$ chains to the cell surface (Hombach et al., 1988) and is probably encoded by the B cell specific gene mb-1, the human homologue of which has an autosomal location (M. Parkar, Institute of Child Health, unpublished observations).

Various murine pre-B cell-specific genes have recently been isolated by subtractive cloning (Yancopoulos et al., 1990). Four have been identified that are expressed in pre-B cell lines but not in B cell lines and several which are expressed in both pre-B cells and B cells. The genomic locations for the genes encoding these polypeptides have not yet been determined. It remains to be seen whether they have human homologues and whether any of them are X-linked.
Understanding of the growth factor requirements of human pre-B cells during differentiation is still hampered by the absence of an in vitro culture system that reproducibly supports B-cell lineage differentiation. Recently, Villablanca et al. (1990) have reported an in vitro culture system which supports the differentiation of normal fetal pre-B cells. These workers cultured CD10+, surface IgM- cells without any feeder cells in RPMI culture medium and 10% fetal calf serum (FCS). After two days, a significant population of cells expressing surface IgM were observed. Southern blot analysis of the kappa L chain locus indicated that rearrangement of this locus had occurred. The reason that this system supported pre-B cell differentiation may have been due to any number of unknown factors in the FCS, however, which would strongly limit the reproducibility of the system.

1.5 Comparison of Human and Murine X-Linked Immunodeficiencies

It is reasonable to assume that a murine X-linked gene homologous to the XLA gene does exist for the following reasons:-
1) There is a considerable degree of similarity within lymphoid development in the murine and human immune systems. Many proteins identified as being expressed during lymphocyte ontogeny in the mouse have already been demonstrated to have human homologues.
2) Genes which are found on the X chromosome of one mammal are found on the X chromosome of all other mammals. This is thought to be linked to X-inactivation, the dosage compensation mechanism which ensures the hemizygous expression of X-linked genes (Ohno, 1969).

One possible murine homologue to a human XID that has already been identified is the scurfy (sf) mouse which may be related to the human WAS. Both are X-linked traits and have similar hematological abnormalities (Lyon et al., 1990).
1.5.1 Candidate Murine Homologues for XLA

Two murine X-linked immunodeficiencies which affect B cells have been identified so far indicating that, as with the human, the murine X chromosome harbours genes which are important in lymphocyte development. These are the DBA/2Ha mice which, in an X-linked manner, transmit B cell unresponsiveness to a T cell replacing factor (Takatsu et al., 1981), and the CBA/N mouse strain which has the X-linked recessive immunodeficiency disorder xid. Male CBA/N mice have a diminished number of B lymphocytes and these have an immature pattern of B lymphocyte surface Ig and lack the differentiation antigens Lyb 3, 5 and 7. They have low serum IgM and IgG3 and they are unable to raise antibodies to polysaccharide antigens (reviewed in Berning et al., 1980). Studies using monoclonal antibodies have identified a B cell surface antigen (BLA-2 or 14G8) that is expressed on most immature B cells but not on a subpopulation of mature splenic B cells. Mice bearing the xid defect lack this population of cells (Hardy et al., 1984). It has been proposed that the XLR gene family may be responsible for the xid defect (Cohen et al., 1985a). This is a series of X-linked DNA segments recognised by the cDNA clone pX310, which was isolated from a lymphocyte-specific cDNA library (Cohen et al., 1985a). This proposal rests upon the basis that firstly, RFLP analysis of the XLR segments in congenic strains of mice with xid indicates that at least one or more of the XLR genes is closely linked to the xid locus. However, recombinations have been found to occur between the locus and some members of the XLR family, indicating that the XLR family is dispersed on the X chromosome (Cohen et al., 1985a). Secondly, XLR expression is lymphoid specific. In the B lineage, it is absent in pre-B or immature B cells (surface IgM+, 14G8+) but present in more mature B cell lines which had demethylated or activated their J chain gene. In contrast, no XLR expression is found in plasmacytomas from xid mice which were expressing J chain (Cohen et al., 1985b).

Probe pX310 was used to screen cDNA libraries for XLR transcripts (Seigel
et al., 1987). One cDNA species designated pM1 was found represent the major transcript found in B and T cells. Probe pM1 was used in a study to construct a genetic map of the murine X chromosome (Mullins et al., 1990). pM1 was found to recognise two discrete loci, designated Xlr-1 and Xlr-2. Both loci were mapped proximally to G6pd (see Figure 1.4) and were therefore found to be distinct from the xid locus, which has been localised distal to tabby (Ta) on the mouse X chromosome and proximal to the hypophosphatemia (Hyp) locus with genetic distances Ta - 6.6 ± 1.8 - xid - 12.2 ± 2.3 -Hyp (Berning et al., 1980).

1.5.2 Comparison of Human and Murine X Chromosomes

Comparative mapping studies of the human and murine X chromosomes have identified five major homologous subchromosomal regions on the two chromosomes (Amar et al., 1988). By examining these regions, it can be seen that the xid defect in the mouse maps to the corresponding region in the human that contains the XLA locus (see Figure 1.4). Overall, there is a strong degree of homology between the two species at these regions. As well as the GLA/Ags and PLP/Plp genes, a number of arbitrary cloned human fragments also recognise murine sequences at this locus, eg. cx52.5 (human = DSX101; murine = DXPas20) and 212/9 (human = DXS178; murine = DXPas15). On positional grounds, therefore, any candidate gene for the xid defect that may be isolated should be considered to be a potential murine homologue of the XLA gene, as should any gene from this region of the murine X chromosome which is expressed in the lymphoid cell lineages. The fact that the phenotype of XLA and xid differ should not lead to the discounting of this possible link. There are precedents for mutations in homologous genes having different phenotypic manifestations even when the expression patterns of these genes may be similar. This may be in part because mice with inherited disorders are inbred and so represent the phenotype of a single mutational event. The human DMD gene (Koenig et al., 1987) and the murine Dmd genes (Bulfield et al., 1984) are good
Figure 1.4
Comparison of the Human and Murine X Chromosomes*

Adapted from Amar et al., 1988. Key: see List of Abbreviations. Loci with different symbols which are homologous are EDA/ Ta; GLA/ Ags; F9/Cf-9 and F8/ Cf-8. Regions on the two chromosomes which share conserved syteny or gene order are similarly shaded. Arrows indicate possible sites of inversions or translocations during evolution of the two chromosomes, as suggested by Amar et al., 1988. These are schematic and are not meant to imply that the murine X chromosome evolved from the human X chromosome or vice versa. The positions of Xlr-1 and Xlr-2 are taken from Mullins et al., 1990.
examples of this. Whereas both species have elevated creatine kinase levels and necrosis of muscle tissue, the DMD patient is usually severely affected by muscle wasting whereas the Dmd mouse is virtually asymptomatic. Thus the Dmd mouse may be more representative of the milder Becker muscular dystrophy, caused by mutations of the same gene, than of DMD.

1.6 Positional Cloning
1.6.1 Introduction

The research described in this study forms part of a long-term positional cloning project to isolate the XLA gene. The general strategy for such approaches has already been described in Section 1.2. Positional cloning tactics have now been used successfully to clone several clinically important genes, including those whose dysfunction cause DMD (Koenig et al., 1987), cystic fibrosis (CF: Rommens et al., 1989a; Riordan et al., 1989), X-CGD (Royer-Pokora et al., 1986), autosomal dominant retinoblastoma (ADRB: Dryja et al., 1990) neurofibromatosis Type 1 (Cawthon et al., 1990; Viskochil et al., 1990) and TCD (Cremers et al., 1990).

In some cases, the establishment of a disease locus has coincided with that of a cloned gene whose expression pattern is consistent with its consideration as a candidate gene. It is then necessary to examine the gene's coding and regulatory sequences for mutations in affected individuals. This was the case, for example, with ADRB (Dryja et al., 1990); for Pelizaeus-Merzbacher disease which is caused by defects in the gene encoding proteolipid protein (PLP) (Willard and Riordan, 1985; Gencic et al., 1989; Hudson et al., 1989; Trofatter et al., 1989) and for Alport syndrome which is caused by mutations in the α5(IV) collagen chain gene (Barker et al., 1990; Zhou et al., 1991).

In the majority of cases, however, there are no candidate genes which have been previously cloned. In the absence of other candidate molecules, it is then necessary to screen the disease locus thoroughly for candidate genes. In the rare instances where large chromosomal deletions are associated with disorders,
"PERT" cloning may be used to construct a library of clones which represent the deleted DNA (e.g. Kunkel et al., 1985). This technique was applied successfully to clone the genes for DMD and X-CGD.

In some instances, there are no gross chromosomal abnormalities associated with a disease and an intensive cloning program is necessary in order to saturate the gene locus. This approach was used successfully to clone CF, however the locus had already been narrowed to a maximum of 1.4-1.9 Mb after extensive linkage analysis and physical mapping studies (Rommens et al., 1989b; Fulton et al., 1989).

Probes which are closely linked to a gene locus can be used to screen libraries of DNA which has been cloned into bacteriophage lambda (Murray and Murray, 1974) cosmid (Collins and Hohn, 1978) or yeast artificial chromosome (YAC) vectors (Burke et al., 1987). Positive colonies can then be screened for conserved and transcribed sequences. The maximum size that can be cloned into lambda and cosmid vectors (< 23 kb and 45 kb, respectively) is restricted by the amount of DNA that can be packaged into bacteriophage heads and this has placed constraints on the usefulness of these vectors. It is possible to undertake chromosome walks, by using the end fragments of cloned DNA segments to rescreen libraries and clone contiguous DNA fragments. In practice, these are extremely labour intensive, and each "step" is unlikely to be greater than 20 kb. One example of this was the analysis of the human major histocompatibility complex class III region, where 61 cosmids were required to cover 560 kb of DNA (Sargent et al., 1989). In addition, there are a number of regions in the genome which are difficult to "walk" past in cosmid or lambda vectors, either because they are highly repetitive or because they are under-represented in such libraries.

The advent of YAC technology has revolutionised cloning of large DNA segments. Inserts of over 1 Mb can be accommodated in these vectors which contain yeast telomeres, although in practice, the average insert size is less than this (e.g. 350 kb; Anand et al., 1990). YACs have now been used successfully
to clone large stretches of DNA, such as the 2 Mb contig around the human Factor IX locus which was produced with 16 YACs (Schlessinger, 1990). Even with the advantages of YACs, it is important to have a clear idea of the size of the disease locus. A genetic distance of 5 cM might in fact relate to a physical distance of 2 Mb or 8 Mb, and whereas it would be feasible to attempt to clone the former, it would be far more advisable to concentrate on generating new polymorphic probes and refining the disease locus in the latter case. One recently developed technique that has made the precise measurement of large physical distances possible is pulsed field gel electrophoresis (PFGE).

1.6.2 Pulsed Field Gel Electrophoresis

This technique facilitates the electrophoretic separation of DNA fragments ranging in size from a few kilobases (kb) to several megabases (Mb). DNA which has been digested by infrequently-cutting restriction enzymes (which are described in detail in Section 1.7.6) can be separated and, following conventional Southern blotting, can be hybridised with DNA probes of interest. In this way it is often possible to link probes which lie hundreds or thousands of kb apart and thus determine the physical distance separating these markers.

Transcribed DNA has few features that are distinguishable from non-coding DNA. One of these features is CpG islands, which are discussed in detail in Section 1.7. A particular advantage of using PFGE for physical mapping is that it facilitates the localisation of CpG islands. This makes it possible to direct cloning attempts at these loci, and once DNA is cloned, to examine these regions for transcribed sequences.

Polyacrylamide and agarose gel electrophoresis have been the conventional methods used for separating DNA fragments on the basis of size. However, the pore sizes of these matrices limit the sizes of the DNA fragments that can be resolved and without significantly reducing the gel concentration, it is not possible to separate fragments of greater than 50 kb (Fangman, 1978). PFGE
does not suffer from this size limitation. The technique utilises at least two sets of electrodes to periodically change the orientation of the electric field in which the DNA is moving.

Each time the direction of the field is altered, the DNA fragments must reorientate before moving towards the new anode(s). The time taken for this reorientation is dependant on the size of the DNA fragment. Smaller molecular weight fragments can change direction more quickly and hence migrate faster in a net forwards direction than can larger molecules.

The technique of PFGE was first described by Schwartz et al. (1983). Their system utilised one inhomogeneous and one homogeneous electric field to separate chromosomes of *S. cerevisiae*. A similar apparatus was described by Carle and Olson (1984a) whose orthogonal field alternation gel electrophoresis (OFAGE) used two inhomogeneous fields. These apparatus, combined with a new technique of DNA preparation, facilitated the separation of intact chromosomes from lower eukaryotes (Carle and Olson, 1984b; Schwartz and Cantor, 1984).

The use of inhomogeneous fields in the early apparatus made the voltage gradient and reorientation angles vary through the gel, making direct size comparisons between adjacent tracks difficult. However, several groups have improved this aspect of the technique by using the combination of homogeneous fields and obtuse reorientation angles. One such system is the field-inversion gel electrophoresis (FIGE) system in which the field is periodically inverted through 180° using a single pair of electrodes (Carle *et al.*, 1986). The switching cycle is timed so that the DNA moves for a longer proportion of time in one direction to produce a net forwards motion and the resulting tracks are straight. Initial observations of anomalously low mobilities of larger DNA fragments with FIGE (Ellis *et al.*, 1987) have been overcome by the application of a "ramped" switching cycle, i.e., one with a gradual increase in both forward and reverse pulse times. A practical disadvantage with FIGE is the tendency for the electrodes to erode prematurely as a result of frequently changing polarity. An
improvement on this system is one in which the gel tray rotates periodically through an obtuse angle and the polarity of the electrodes remains unchanged (Anand, 1986). Chu et al. (1986) applied the principles of electrostatics to produce a contour-clamped homogeneous electric field (CHEF), which uses multiple electrodes in a hexagonal configuration with an obtuse reorientation angle of 120°. The multiple electrodes provide a good working approximation of "infinitely long" electrodes, so the voltage potential along the "length" of the electrode remains even (Figure 1.5). These fields give straight, comparable tracks and good resolution of large chromosomal DNA fragments has been obtained (Vollrath and Davis, 1987). The transverse alternating field (TAFE) apparatus (Gardiner and Patterson, 1988) uses a vertical electrophoresis apparatus with the electric fields being applied at angles through the thickness of the gel. These fields are also homogeneous.

The factors that affect separation of DNA fragments by PFGE are switch time, voltage gradient, temperature, and agarose concentration. The size of the fragments separated increases with the length of the pulse-time.

For each set of running conditions used, there is a window of optimal resolution and a zone of limiting mobility (LM) which defines a fragment size above which the fragments are not resolvable. For a single set of running conditions, the distance migrated by DNA molecules in general bears a linear relationship to their size. This is not strictly correct, however. In practice (using a CHEF apparatus) there often appears to be a small zone of increased resolution approximately one-third of the way down the gel. In addition, small molecules which are of the size that would be separated by conventional electrophoresis in such a matrix will still produce a standard non-linear pattern of separation.

For larger DNA fragments of a size that are normally trapped by the agarose matrix used, there is a linear relationship between $\log_{10}$ switch time and $\log_{10}$ size of fragment separated. Thus, once a set of running conditions has been
A and B represent two sets of electrodes. The arrows indicate the direction of electron flow in the two electric fields, respectively.
established, this relationship can be used to calculate the switch time required to separate a different size range for a fixed voltage gradient. A second useful rule is that to resolve a single fragment size, the product of the switch time and the voltage should remain constant. Typically, to separate smaller DNA molecules (1.5 Mb or less), voltage gradients of 6V/cm or greater can be used. However, for the separation of larger fragments it seems to be necessary to use much smaller voltage gradients in the region of 1.3V/cm. The pulse-time must therefore be increased accordingly.

1.6.3 Radiation Hybrid Mapping

A second technique which is increasingly used in parallel with PFGE to map large regions of DNA is radiation hybrid (RH) mapping. Used alone, this technique does not provide data on physical distances between probes, but it can be used to order probes extremely effectively when sufficient numbers of clones are used. Unlike PFGE, radiation hybrid mapping is not dependant upon the frequently uneven distribution of rare cutting restriction enzyme sites in the genome. The method is analogous to genetic linkage mapping. Lethally irradiated human chromosomes are fused with rodent recipient cell lines, resulting in the retention of human chromosome fragments in the recipient cells. The likelihood that two particular loci will be co-inherited (ie. both retained or both lost) increases with their original proximity on the chromosome. Statistical analysis can then be used to determine the most likely order of small groups of probes. In addition, RH mapping can enable PFGE to be used in a directed manner to physically link probes which have already been shown to be closely linked.

Two methods of generating radiation hybrids have been developed: these are chromosome-mediated gene transfer (CMGT: McBride and Ozer, 1973) and irradiation and fusion gene transfer (IFGT: Goss and Harris, 1975). Recently, RH mapping has been used to create a 20 Mb map of the proximal long arm of chromosome 21 (Cox et al., 1990). In this case, the IFGT technique was used
to generate hybrids using a somatic cell hybrid which retained a single chromosome 21 on a rodent background as the donor cell line. The complimentary use of PFGE also made it possible to equate breakage frequency with physical distance (Burmeister et al., 1991).

A similar adaption of the IFGT technique has been used to generate a panel of approximately 200 hybrids containing fragments from the human X chromosome (Benham et al., 1988). By using a somatic cell hybrid containing a single X chromosome as the donor cells, the only human material in the resulting hybrids is of X chromosome origin.

The IFGT technique has been the preferred method of making hybrids for mapping because the non-selectable nature of fragment retention results in a low frequency of interstitial rearrangements and deletions post-fusion. For example, the analysis of seven clones from the X chromosome series of IFGT hybrids with sets of closely linked probes revealed that all seven hybrids maintained the linear integrity of their insertions (Benham et al., 1989). This contrasted strongly with the findings of Porteous et al. (1986) when CMGT clones selected for HRAS-1 all showed detectable interstitial deletions within a few Mb.

1.7 CpG Islands, Genes and Methylation

1.7.1 The Base Composition of Vertebrate Genomes is not Uniform

Vertebrate DNA is A and T-rich and the dinucleotide 5'-CpG-3' (CpG) occurs at only 0.2 to 0.25 of the frequency expected from the base composition (Josse et al., 1961; Swartz et al., 1962). In addition, between 60% and 90% of CpGs are methylated at the 5 position of cytosine. There is, however a fraction of the genome which is rich in G + C and contains CpG at the expected frequency (Tykocinski and Max, 1984; Bird et al., 1985). In addition this fraction is hypomethylated and can be extensively digested by the methylcytosine-sensitive restriction enzyme Hpa II (Cooper et al., 1983). Frequently called the HTF (Hpa II tiny fragment) fraction, it accounts for approximately 1% of the genome and is dispersed in "CpG islands" averaging 1
to 2 kb in length. There are estimated to be of the order of 30,000 CpG islands in the murine haploid genome with an average spacing of 100 kb, although there is considerable variation about this mean.

1.7.2 CpG Islands are Usually Associated with Genes

The analysis of randomly-picked murine HTF clones by Bird et al. (1985) provided evidence that there was a close relationship between CpG islands and genes. Out of 13 clones picked at random, four were of mitochondrial or ribosomal origin, the latter of which is known to be rich in C+G content. Eight of the remaining clones hybridised to single bands and at intensities which were equivalent to one or a few copies per genome. Three of these were subcloned and hybridised to poly(A)+ RNA from mouse and two out of three detected discrete transcripts in a wide range of murine tissues (Lavia et al., 1987). Lindsay and Bird (1987) picked clones containing a restriction site for the enzyme Sac II, whose recognition sequence, CCGCGG, is rare outside CpG islands. Three out of four clones recognised a transcript when hybridised to poly(A)+ RNA. Toniolo et al. (1988) surveyed a 100 kb stretch of DNA around the human G6PD gene. A total of three CpG islands were found in this region and all were closely linked to genes. This reflects a strong coincident occurrence of genes where there are CpG islands.

Several reports have documented genes that have unmethylated CpG-rich regions associated with their 5' ends. These include the chinese hamster adenine phosphoribosyl transferase (aprt) gene, the mouse dihydrofolate reductase (dhfr) gene (Stein et al., 1983) and the chicken α2(1) collagen gene (McKeon et al., 1982).

1.7.3 Which Genes have CpG Islands?

Gardiner-Garden and Frommer (1987) undertook a study of nearly 200 gene sequences in which total G+C and the ratio of observed/expected CpG were
calculated in a 100 bp moving average window. CpG islands were defined as stretches of DNA where both the moving average of %G+C was greater than 50 and the ratio of observed/expected CpG was greater than 0.6. They observed that all housekeeping genes examined (including those encoding metabolic enzymes, structural proteins and snRNAs) had an associated CpG island that began upstream of the translation start site (termed 5' CpG islands). Other genes with 5' CpG islands included widely (but not ubiquitously) expressed genes and genes whose expression was highly tissue-specific. Other tissue-specific genes either had islands situated 3' of the translation start site (termed 3' islands) or else had no associated CpG island. No difference was obvious between the islands of housekeeping or tissue-specific genes with regard to length or CpG content.

1.7.4 What is the Function of CpG Islands?

As has been stressed before, one of the properties of CpG islands is hypomethylation. There are several reports that indicate that cytosine methylation may play a role in the regulation of gene expression. Bird (1986) postulated that CpG islands may serve to identify sequences that are to be constantly available in the nucleus. However, Beggs and Migeon (1989) found a negative correlation between CpG clusters and DNA regions which were associated with the nuclear scaffold (Scaffold Associated Regions; SARs). In their studies, no SARs were detected in six X-linked housekeeping genes with CpG islands whereas two X-linked tissue specific genes without associated 5' CpG islands both had SARs. As several tissue-specific genes do not have associated islands this would rather negate a universal role of CpG islands in maintaining the availability of such genes for transcription.

There are cases where tissue-specific genes have CpG islands whose methylation state has a close correlation with gene expression, however. For example, Albini et al. (1990) studied the interphotoreceptor retinoid-binding protein (IRBP) whose expression is limited to retinal photoreceptor cells and a
subset of pinealocytes. The first exon and the 5' promotor region both contain CpG-rich regions which do not fall into the classical definition of CpG islands, ie. observed CpG content is 36% and 32% of expected, respectively. Both of these regions were observed to be hypomethylated in IRPB-expressing tissues and methylated in non-expressing cells. Interestingly, a Hha I site (CCGG) between these two regions was methylated in both cell types.

1.7.5 CpG Islands on the X Chromosome

On the autosomes, the CpG islands associated with housekeeping genes are hypomethylated. The investigations of several workers into the methylation status of cytosine residues situated at various locations (within introns, exons, or CpG islands) along the X chromosome has shown that this is not the case with the X chromosome. It has been proposed that methylation plays a special role in the maintenance of transcriptional silence of genes on the inactive X chromosome. Toniolo et al. (1988) investigated the correlation between expression of the G6PD gene and the methylation status of three CpG islands within a 100 kb region on the X chromosome. One island was associated with the 5' region of the G6PD gene; the other two were positioned 3' of the G6PD gene and were at the 5' ends of the nearby genes GdX and P3. G6PD is constitutively expressed and both GdX and P3 show expression in many tissues. All three of these islands were found to be methylated on the inactive X chromosome and unmethylated on the active X which expressed G6PD. In addition, the reactivation of G6PD expression in cell lines correlated with coincident demethylation of all three islands. Toniolo et al. (1988) proposed that methylation of CpG islands was important in maintaining regional inactivation of the X chromosome. This relationship only appeared to be applicable with selected CpG sites. These workers found no difference in the methylation status of CpG sites that were not contained in islands. In addition, they were unable to find any non-gene associated islands within a 100 kb region and proposed that
the majority of CpG islands on the X chromosome would be associated with
genes. Lindsay et al. (1985) looked at the methylation status of CCGG sites in
four X chromosome clones. No correlation was found between the methylation
status of these sites and the activation state of the X chromosome. Importantly,
only one of these clones, pGK824 was associated with a transcribed sequence
and this particular clone excluded the 5' end of the gene. PGK was later found
to have a 5' CpG island containing CpG sites whose methylation status
correlated strictly with X inactivation (Keith et al., 1986). Wolf et al. (1984a)
observed a similar phenomenon associated with the X-linked gene hypoxanthine
phosphoribosyl transferase (HPRT). Hypomethylation of the 5' end of the gene
was shown on the active X chromosome compared to the inactive X
chromosome. In contrast, the body of the gene was highly methylated on the
active X whereas the inactive X was variably methylated. There are examples of
a reversal of this phenomenon, for example the X-linked DNA marker DXS255
(M278), which is not a gene sequence, is used to differentiate between active
and inactive X chromosomes on the basis that it detects hypomethylated CpG on
the inactive X chromosome (Boyd and Fraser, 1990). The inactive and active X
chromosomes do not appear, therefore, to be differentially methylated along
their entire length. However there are specific CpG dinucleotides which are
situated within gene-associated CpG islands for which hypomethylation and
activation status (and therefore potential for gene expression) are correlated.

All the genes discussed so far in this section have been of the housekeeping
variety and no mention has been made of genes with limited tissue expression.
On the autosomes, all housekeeping genes have CpG islands but a proportion of
tissue-specific genes do not have associated islands. Two X-linked genes with
tissue-specific expression are blood clotting Factor IX (Anson et al., 1984,
Salier et al., 1990) and ornithine transcarbamylase (OTC) (Davies et al., 1985).
Neither of these genes have 5' associated CpG islands. Physical mapping
around the steroid sulphatase (STS) locus at the distal short arm of the human X
chromosome has revealed a cluster of CpG rich sites 5' of the gene locus (Li et
al., 1990). Few tissue-specific genes that map to the X chromosome have actually been examined for the presence of CpG islands.

Whilst no good correlation has been found between tissue-specific genes and 5' CpG islands on the X chromosome so far, there is ample evidence to suggest that CpG islands do occur frequently in regions of DNA which contain a relatively large proportion of transcriptionally active sequences. For example, although Factor IX is not associated with a CpG island, it is situated in a gene-rich area containing three CpG islands (Nguyen et al., 1987). A pulsed field map of the DMD locus (Burmeister et al., 1988) revealed a clustering of CpG sites spread over several hundred kb near the 5' end of the DMD gene. This coincided with a small G-light band (Xp21.2) within the large G-dark band (Xp21) in which the bulk of the gene is localised. CpG islands appear to correlate with the G-light staining (or Reverse) chromosome bands which are richer in G+C and contain a high gene density, mainly housekeeping and CpG island-containing genes (Bernardi et al., 1985). Physical mapping studies of human chromosome 21 led Gardiner et al. (1990) to propose that the restriction enzymes used in PFGE studies (see below) only recognised a subset of CpG sites which were originally defined using enzymes which cut far more frequently in the genome. They suggested that the CpG islands identified by rare-cutting restriction enzymes might therefore define gene clusters rather than individual genes.

1.7.6 CpG Islands Contain Sites for Infrequently Cutting Restriction Enzymes.

A number of type 2 restriction enzymes cut infrequently in the genome due to the fact that their have hexamer or octamer recognition sites containing one or more CpG dinucleotide. In addition these enzymes are methylation sensitive and will not digest DNA which has methylcytosine in its sequence. The cutting sites for these enzymes are therefore limited to a greater or lesser extent to CpG
islands.

The use of these enzymes in connection with PFGE is thus twofold. As well as producing reproducible large sized fragments in order to link DNA loci, the cutting sites of these restriction enzymes used in combination can indicate the presence of CpG islands. As discussed above (Section 1.7.3), CpG islands indicate the presence of genes which can be of the housekeeping or tissue-specific variety. The frequency of restriction sites within CpG islands is listed in Table 1.1 for a number of enzymes. The majority of sites listed as available in bulk DNA will actually be methylated and therefore cannot be digested by these enzymes.

1.8 Aims of this Study

The research described in this thesis forms part of a long-term attempt to isolate the XLA gene using a "positional cloning" approach, in the absence of candidate genes. The primary aim was to produce a physical map of the XLA locus in the hope that this would incorporate flanking markers and therefore give a maximum physical estimate for the possible locations for the XLA gene between these markers. This would determine the nature of further work, ie. whether to concentrate on linkage data or whether cloning the DNA between the flanking markers was feasible.

A second area of research was developed as the PFGE work progressed. The detection of CpG-rich clusters close to DXS178, the probe most tightly linked to XLA, hindered the construction of a PFGE map and resulted in the commencement of a chromosome walk towards these clusters by screening an X chromosome cosmid library with DXS178.
Table 1.1
Frequencies of Restriction Sites for Enzymes in Bulk DNA and in CpG Islands

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>No. of CpGs in restriction site</th>
<th>Restriction Site</th>
<th>Frequency ** of sites in Bulk DNA (1/ x kb)</th>
<th>No. of sites/kb of CpG island observed*</th>
<th>Percentage of Sites in CpG Islands</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not I</td>
<td>2</td>
<td>GCGGCCGC</td>
<td>6000</td>
<td>0.3</td>
<td>95</td>
</tr>
<tr>
<td>Eag I</td>
<td>2</td>
<td>CGGCCG</td>
<td>250</td>
<td>1.1</td>
<td>73</td>
</tr>
<tr>
<td>BssH II</td>
<td>2</td>
<td>GCGCGC</td>
<td>250</td>
<td>1.2</td>
<td>75</td>
</tr>
<tr>
<td>Sac II</td>
<td>2</td>
<td>CCGCGG</td>
<td>250</td>
<td>1.4</td>
<td>76</td>
</tr>
<tr>
<td>Mlu I</td>
<td>2</td>
<td>ACGCGT</td>
<td>68</td>
<td>0.03</td>
<td>2</td>
</tr>
<tr>
<td>Nru I</td>
<td>2</td>
<td>TCGCGA</td>
<td>68</td>
<td>0.1</td>
<td>6</td>
</tr>
<tr>
<td>Nae I</td>
<td>1</td>
<td>GCCGGC</td>
<td>62</td>
<td>0.9</td>
<td>36</td>
</tr>
<tr>
<td>Nar I</td>
<td>1</td>
<td>GGCGCC</td>
<td>62</td>
<td>1.2</td>
<td>42</td>
</tr>
<tr>
<td>Sfi I</td>
<td>0</td>
<td>CCCCCNNNNGGGGG</td>
<td>390</td>
<td>0.15**</td>
<td>37</td>
</tr>
<tr>
<td>Xho I*</td>
<td>1</td>
<td>CTCGAG</td>
<td>23</td>
<td>0.3**</td>
<td>6</td>
</tr>
</tbody>
</table>

*Values taken from Bird, 1989.
** Estimated values. Bulk DNA was taken to comprise 60% A+T with the CpG dinucleotide occurring at 25% of the expected frequency. CpG islands were taken to comprise 65% C+G with the CpG dinucleotide occurring at the expected frequency. The percentage of sites in islands was calculated using the observed frequency of sites in islands and the calculated frequency of sites in bulk genomic DNA. The size of the genome was taken to be 3 x 10^9 bp and the total number of CpG island to be 30,000.

* Xho I contains A+T in the recognition sequence and therefore may occur less frequently in CpG islands, as has been observed for Mlu I and Nru I.
1.8.1 Rationale for the Choice of Probes used in this Study

At the beginning of this study in September, 1987, genetic studies had shown that the XLA locus showed linkage to the probes DXS3, DXS94 and DXS17 in proximal Xq. No recombinations had been found between XLA and the latter two probes. Unpublished data from this Department also indicated that the probe DXS178 was linked to the XLA locus with no recombinations seen in ten meioses. In addition, the probe DXS3 had been found to be deleted in a choroideremia patient MBU, who had a deletion within Xq13-2.13 (Hodgson et al., 1987). This patient did not have XLA. It was therefore decided that it was reasonable to attempt to construct a physical map encompassing the MBU deletion breakpoint and the loci which showed no recombination with XLA. Such a map would be extremely useful in pinpointing the XLA locus as more genetic data emerged. Arveiler et al. (1987) published a genetic study which indicated that the order of these loci was centromere - DXS3 - (DXS178, DXS94) - DXS17 - telomere.

In 1987, the markers which showed linkage to XLA were broadly localised in Xq21.3-Xq22, with only DXS94 being finely localised to Xq22 (see Figure 1.6). At this point in time, it was quite possible that the markers used in the linkage studies would be situated close to each other and so I commenced PFG studies with DXS17, DXS94 and DXS178. The additional probes DXS87 and DXS88 were also employed and I began to obtain additional markers from Xq21.3-Xq22, in order to "saturate" the region and increase the possibility of physically mapping the polymorphic probes linked to XLA.

By mid 1988, the only probes successfully linked were DXS87 and DXS88, which later were found to recognise the same Bgl II RFLP. In the same year, Cremers et al. published an article which for the first time ordered a large number of DNA loci in proximal Xq by scoring their copy number in DNA from a series of choroideremia patients who had large deletions or duplications in this region (see Figure 1.6). These loci included a number of markers which had
Figure 1.6
Localisation of Probes Used in this Study

Kwan et al., 1990

Cremers et al., 1988

Davies et al., 1987

<table>
<thead>
<tr>
<th>21.1</th>
<th>21.2</th>
<th>21.31</th>
<th>21.32</th>
<th>21.33</th>
<th>22.1</th>
<th>22.2</th>
<th>22.3</th>
<th>23</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xq</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
previously been undesignated or loosely localised, including PLP, DXS24 and DXS211 and this presented me with an opportunity to acquire a number of additional probes from the region of interest i.e. in Xq22, distal to the choroideremia deletions. The probes obtained were DXS101, PLP, DXS24, DXS147 and DXS211. The sources of these probes are listed in Appendix A1. In addition, I obtained probes DXS83 and DXS114 whose designations had not been refined from Xq21.3-Xq22 (see Figure 1.6).

DXS101 actually hybridises to a small family of sequences, all mapping to the Xq22 region and it would have been possible to generate more single-copy probes from the region by obtaining DXS101-positive cosmids and isolating single-copy sequences. However, in view of the large number of additional markers I had obtained, I chose to concentrate on using these other markers first.

In 1990, Kwan et al. reported that recombinations had been found between DXS94 and DXS17. This indicated firstly that these markers flanked the XLA locus on the telomeric side, and secondly that they were further from DXS178 and XLA than previously suggested. The knowledge that there were now markers which flanked the XLA locus both proximally and distally meant that the inclusion of such loci on a contiguous PFG map would therefore encompass the XLA locus and give a maximum value to the size of DNA in which it was situated.
All reagents were purchased from Sigma unless otherwise stated.

2.1 Tissue Culture

2.1.1 Lymphoblastoid Cell Lines

Lymphoblastoid cell lines (MBU, NP, DM and LCL-127) were cultured in RPMI 1640 (Gibco) with 10% fetal calf serum (FCS), 2mM glutamine and 40μg/ml gentamycin (Cidomycin, Roussel).

2.1.2 Fibroblast Cell Lines

Fibroblast lines (AnLy and RAG) were cultured in the same medium used for lymphoblastoid cell lines. Between 3-5mls Trypsin-EDTA (Gibco) was used to loosen cells prior to feeding. AnLy is a somatic cell hybrid containing a derivative X chromosome (Xq1.2-ter) from an anhydrotic ectodermal dysplasia (EDA) patient with a balanced X/9 translocation. RAG is a murine fibroblast line and is the rodent parent for AnLy.

2.1.3 X Chromosome Radiation Hybrid Lines

The X chromosome radiation hybrid lines were cultured in DMEM (Gibco) with 10% FCS, 2mM glutamine, 40μg/ml gentamycin and HT supplement (Gibco) by Ms. Karen Curtiss (Institute of Child Health, London).

2.1.4 Preparation of Phytohemagglutin (PHA)-Stimulated T Cell Blasts

The PHA-stimulated T-cell blasts were prepared by Mr. Michel de Weers (Leiden University Hospital, The Netherlands) as follows. Patient and control peripheral blood samples were taken into preservative-free heparin and separated on Ficoll gradients (see Section 2.2.1). 2x10^6 Patient cells were mixed with 10^7 control cells which had been irradiated at 2000 Rads. The cells were cultured at
a concentration of $1 \times 10^6$ cell/ml in Iscove's Modified Dulbecco's medium (Gibco) with 20 units/ml penicillin and streptomycin, 100 units/ml interleukin 2 (IL2), 0.5% PHA and 10% human AB serum. The cultures were split and fed every three days with the same medium. After 8 days the cells were harvested and DNA was prepared as described in Section 2.2.3.

2.2 Preparation of DNA in Agarose Blocks for PFGE

2.2.1. From Blood

The methods used for agarose block formation, enzyme digestion and blotting for PFGE were adapted from a protocol supplied by Drs. Barlow and Lehrach (ICRF, London). 10-60mls of blood, taken into 100µl preservative-free heparin, were mixed with Dextraven-110 (Fisons plc) at a ratio of 2:1. The mixture was left to sediment at room temperature for 45 minutes, after which the leukocyte-rich plasma was removed into a fresh tube and washed twice with phosphate buffered saline (PBS) "A". The cells were resuspended in 1-6mls PBS "A" and the cell concentration was determined by counting a 1/10 dilution in 1% (v/v) acetic acid using a hemocytometer. The cell concentration was adjusted to between 14-22x10^6 cells per ml which produced agarose blocks containing between 0.5-0.75x10^6 cells.

Several block formers (LKB) were sealed with autoclave tape and placed on a clean sheet of glass on a bed of ice. A 1% (w/v) solution of low gelling temperature agarose (InCert, FMC) was prepared and held at 37°C. The cell suspension was mixed in 500µl aliquots with an equal volume of agarose and 70µl aliquots were then dispensed into the block formers. After 40 minutes, the blocks were tipped out onto a clean piece of plastic film and placed 30 at a time into 50ml polypropylene tubes. To each tube 2.5mls of protein digestion buffer was added (0.45M EDTA, 1% sodium dodecyl sulphate (SDS), 2mg/ml proteinase K). The tubes were incubated for 40-60 hours at 50°C. The digestion buffer was then removed by four 50ml rinses in TE pH 8.0 (10mM Tris-HCl, 1mM EDTA). Finally, to remove any residual proteinase K and SDS, the
blocks were incubated for 30 minutes at 55°C in TE (pH 8.0) containing 40μg/ml phenylmethylsulphonylfluoride (PMSF). The PMSF stock was freshly prepared by dissolving 40mg PMSF in 1ml isopropanol at 65°C. The PMSF incubation step was repeated and then the blocks were ready to be digested with a restriction enzyme or stored at 4°C.

2.2.2. From Sperm

A fresh ejaculate of sperm cells was washed twice in PBS "A" and resuspended in 2-3mls PBS. The cell concentration was determined using a hemocytometer and the concentration was adjusted to between 28-44x10⁶/ml (twice the concentration of the leukocytes because sperm are haploid). The same protocol was then followed as for the blood blocks except that the protein digestion buffer additionally contained 10mM dithiothreitol (DTT) to remove protamines tightly bound to the DNA (Burmeister et al., 1988).

2.2.3. From Cell-Lines

Lymphoblastoid and fibroblast cell-lines were harvested 24-48 hours after addition of fresh medium so that the majority of cells were not dividing rapidly. Adherent cells were trypsinised and trypan blue was used to determine viability. Cultures that were less than 80 percent viable were separated on Ficoll gradients to remove dead cells as follows: cells were layered carefully onto an equal volume of Ficoll-Paque (Pharmacia) and centrifuged at 550g for 20 minutes at room temperature. The cell-rich interface was removed with a pipette into a clean flask and washed twice in PBS. The cell pellet was resuspended in PBS and the concentration adjusted to between 14-22x10⁶ cells/ml. The same protocol was then followed as for the blood blocks.

2.2.4 Storage of DNA Blocks

Blocks were stored at 4°C in a solution of 0.5M EDTA (pH8.0). No loss in
DNA quality was found over several months of storage. The EDTA solution was replaced every 3 months.

2.2.5 Assessment of DNA Blocks for Degradation

Prior to enzyme digestion, blocks were assessed for DNA degradation. One block from each new batch was incubated overnight at 37°C in 100μl H₂O. This block was then loaded onto a gel and electrophoresed using PFGE with conditions designed to separate up to 1 Mb DNA. If the blocks were of good quality, no significant DNA migration was seen out of the wells. If degradation was apparent, the protein digestion step was repeated for 20 hours and the blocks retested.

2.3 Enzyme Digestion of DNA in Agarose Blocks

Blocks that had been stored in EDTA solution were rinsed in 200mls of TE (pH 8.0) three times for 20 minutes. To a 1.5ml tube the following was added: 100μl H₂O, 10μl bovine serum albumin (stock 10mg/ml), 20μl 10X enzyme buffer, 10μl spermidine (stock 0.1M) and finally 30 units restriction enzyme (New England Biolabs). The digestion mixture was vortexed and centrifuged briefly, after which a 70μl DNA block was added. The tube was incubated at the appropriate digestion temperature for 16 hours. Spermidine was omitted from enzyme digests with low salt (<50mM NaCl) requirements.

For sequential enzyme digests, the DNA was first incubated with the enzyme with the lowest salt requirement as described above. The DNA block was then tipped into a clean weighing boat and the digest mixture removed carefully with a pipette. The block was then rinsed three times for 20 minutes each in TE (pH 8.0) and the second enzyme digest set up as described above. The second digest was carried out for 5-6 hours and the blocks were then placed at 4°C for one hour prior to loading into the gel.
2.4 Preparation of Size Standards for PFGE

Two types of size markers were used to standardise PF gels. For the majority of runs, whole yeast chromosome preparations were used. For runs when fragments of under 200 kb were resolved, bacteriophage lambda concatamers were used. The sizes of the various markers are as follows. *S. cerevisiae* strain YP148 was provided by Dr. Don Williamson, (NIMR, Mill Hill, London) and the size markers are 92, 213, 276, 351, 441, 550, 598, 681, 752, 791, 824, 940, 970, 1025, 1115, 1500, and 2500kb. *H. wingei* was provided by Dr. Peter Jones (MRC, Cambridge) and the sizes are 1.03, 1.25, 1.5, 1.8, 2.6, 2.9 and 3.3Mb. *S. pombe* sizes are 3.5, 4.5 and 5.7 Mb. Yeast chromosome sizes were taken from Jones *et al.*, 1989. Bacteriophage lambda is approximately 48.5kb and concatamers were multiples of this molecule.

2.4.1 Preparation of Yeast Chromosomes for PFGE Size Standards *(Jones et al., 1989.)*

10mls of YEPD medium (see Section 2.18) were inoculated with a single colony of a desired yeast strain and incubated at 200 rpm for 16 hours at 30°C. This inoculate was then added to a further 90mls of YEPD and incubated as above until the rate of yeast division had reached late log phase, equal to a concentration of 2.5x10⁸/ml. This took approximately 15 hours for *S. cerevisiae* or *H. wingei* and 40 hours for *S. pombe*. The yeast concentration was determined using a hemocytometer.

The yeast were divided into two 50ml tubes and pelleted at 1200g for 3 minutes. The supernatant was discarded and the pellets were resuspended vigorously. The tubes were then filled up with 50mM EDTA (pH 7.5) and the yeast were pelleted as before. The EDTA solution was discarded and the pellet was resuspended.

A solution of 1% (w/v) LMP agarose (SeaPlaque, FMC) was made in 50mM EDTA (pH7.5) and held at 42°C. 40mg Zymolyase-20T (Sigma) was added to
6mls CPES solution (120mM Na₂HPO₄, pH 6.0 containing 40mM citric acid, 1.2M sorbitol, 5mM DTT, 20mM EDTA pH7.5). The solution was agitated gently to facilitate solubilisation of the zymolyase and the solution was then passed through a 0.2μ filter. The yeast were placed in a 42°C waterbath for two minutes. To each 50ml tube was added 3mls CPES solution with zymolyase and 5mls LMP agarose. After mixing, the contents were pipetted in 100μl aliquots into block formers at room temperature. Once the blocks had set, they were tapped out onto plastic film and incubated at 30°C in 200ml CPE solution (120mM Na₂HPO₄ pH6.0 containing 40mM citric acid, 20 mM EDTA) for 3 hours. The CPE solution was then removed and the plugs were incubated in 20mls solution 3 (10mM Tris-HCl pH8.0 containing 0.45mM EDTA, 1% SDS, 1mg/ml proteinase K) at 50°C overnight. The blocks were then ready for use and were stored in 0.5M EDTA (pH 8.0) at 4°C until required. Between 1/3 and 1/2 of a block was used per track.

2.4.2 Long Term Storage of Yeast

Yeast colonies were stored on YEPD plates for several months at 4°C. Fresh plates were made every 2 months for H. Wingei and every 6 months for the other two strains. Stabs were also kept on Sabarose slopes at 4°C. Overnight inoculants were stored with 10% glycerol in 1.5ml freezing vials at -70°C.

2.4.3 Preparation of Lambda Concatamers (Anand, 1987).

Lambda DNA which had not been previously frozen was mixed at a final concentration of 200μg/ml with 3% Ficoll Paque (5.7% Ficoll, 9% diatrizoate disodium, Pharmacia) and 2xSSC (see Section 2.18). The mixture was incubated at 37°C for 30 minutes and then at room temperature for 16 hours. 3-5μl was sufficient for each gel track. The lambda DNA was handled at all times with 1ml Gilson pipette tips which had been widened using a scalpel to minimise shear forces.
2.2.5 Pulsed Field Gel Electrophoresis

All gels were run using LKB Pharmacia Pulsaphor apparatus. The running conditions used are described in the legend for each figure in the Results sections. The samples were loaded onto a 170ml agarose (SeaKem GTG, FMC) gel using a scalpel and a 1ml pipette tip. The running buffer used was 2.5l 0.5 x TBE (pH 8.0) and the temperature of the buffer was controlled using a thermostatic circulator (LKB MultiTemp II or Techne RB 5) so that the final temperature of the buffer was between 10-14°C, depending on the run conditions of the gel. The buffer was precooled for 2 hours prior to electrophoresis.

The LKB hexagonal electrode kit was used. This is a CHEF system, which is described in Section 1.6.2. The diameter of the CHEF apparatus was 28.5cm. The switching of the power supply between sets of electrodes was controlled using either an LKB Pulsaphor apparatus or a Flowgen switch unit. Gels that were run for long periods had the buffer replaced at least every three days.

2.6 Southern Blotting onto Nylon Filters

After electrophoresis, the gels were stained for 20 minutes in 300mls of 0.5 μg/ml ethidium bromide and photographed. The gels were then depurinated for 20 minutes in 240mM HCl, rinsed in denaturing solution (1.5M NaCl, 0.5M NaOH) and soaked in denaturing solution for two 30 minute periods. Afterwards, the gels were rinsed in neutralising solution and soaked in neutralising solution for two 30 minute periods. The gels were blotted onto either Hybond-N or Hybond-N+ filters (Amersham) according to the manufacturer's instructions, using a foam platform covered in Whatman 3MM Chromatography paper and soaked in 20xSSC. A stack of tissues was placed on top of the blot and the gels were blotted for 24-40 hours. After the first 16 hours, the wet tissues were replaced and the 20xSSC levels were replenished. Following blotting, the gels were fixed appropriately: the Hybond-N filters were rinsed in 2xSSC and baked in a 80°C oven for 2 hours whereas the Hybond-N+
filters were placed in 0.4M NaOH for 60 minutes, rinsed twice in 2xSSC and air-dried.

2.7 Prehybridising and Hybridising the Filters

2.7.1 Prehybridisation

The filters were dampened in 2xSSC and sealed in a plastic bag, leaving a hole at one corner. The prehybridisation solution was composed of 10x Denhardts solution (see Section 2.18), 6xSSC, 10% (w/v) dextran sulphate, 1% SDS and 5μg/ml sonicated salmon sperm. The salmon sperm was boiled for 5 minutes immediately before addition. Frequently, a precipitate formed when the prehybridising solution was made. This was resolved by warming the solution for a few minutes at 65°C. For standard 15 x 15 cm filters, 10mls of prehybridisation solution was used. The solution was pipetted into the bag, air bubbles were squeezed out and the bag was sealed. The filter was incubated in a 65°C waterbath for a minimum of three hours.

2.7.2 Hybridisation

After this time, [32P]dCTP labelled DNA was denatured by boiling for five minutes and pipetted into the bag at a concentration of 10⁶ dpm per ml of prehybridisation solution. If only a few μls of probe was required, the required aliquot was added to 1ml of prehybridisation solution (without dextran sulphate) and then added to the bag to aid mixing. The bag was then resealed and rubbed thoroughly with a wad of tissues to distribute the probe. The filters were incubated for a further 16-20 hours at 65°C.

2.7.3 Removal of Non-Specifically Bound Probe

Filters were rinsed in the following solutions to remove any probe which had bound non-specifically. First, they were washed in a solution of 3xSSC and 0.1% SDS for 20 minutes at room temperature. This step was repeated twice. The filters were then washed for 30 minutes at 65°C in a solution containing 0.1%
SDS and either 3xSSC, 1xSSC, 0.5xSSC or 0.2xSSC, depending on the stringency requirements of the particular probe. The filters were monitored after washing and were deemed to be sufficiently washed if the space above the wells read no higher than background radiation using a hand-held monitor. The filters were placed on 3MM paper to remove excess surface moisture.

2.7.4 Autoradiography

The filters were then wrapped in plastic film and exposed to X-ray film (Kodak X-AR5) with two intensifying screens (Lightning Plus, Cronex, Dupont) for 1-3 days. The films were developed and the filters were re-exposed for a further time period if required.

2.7.5 Stripping Filters

In order to re-hybridise filters, bound probe was removed by pouring 500ml of boiling 10mM Tris-HCl (pH 8.0) onto the filters. The filters were left shaking in the solution on a rotating platform and allowed to cool until the temperature of the solution reached room temperature. In order to check that the probe had been removed the filters were exposed to X-ray film as before.

2.8 Transformation of *E.Coli*

DNA probes were frequently received in the form of plasmid DNA in aqueous solution or as an ethanol precipitate. In order to grow up large quantities of DNA the plasmids were introduced into *E.Coli* (strain HB101) as follows.

2.8.1 Preparation of Competent Cells

HB101 from a frozen glycerol stock was streaked onto a petri dish containing LB-agar and incubated overnight. The following day, a single colony was inoculated into 10mls LB medium and incubated overnight. On day three,
1.5mls of the inoculant was added to 50mls LB medium and incubated for 1 - 1.5 hours until the optical density of the medium reached 0.44 absorbance units at 550nm. The *E. Coli* were pelleted by centrifugation at 1100g for 10 minutes at 4°C. The supernatant was discarded and the cells were resuspended in 10mls of 10mM NaCl. The cells were pelleted as before and resuspended in 20mls of 100mM CaCl₂ and left on ice for 20 minutes. The cells were again pelleted. This time they were resuspended in 4mls 100mM CaCl₂ and left on ice for 15 minutes.

### 2.8.2 Transformation of Competent Cells

300μl of competent cell suspension were mixed with 100-200ng of the desired plasmid DNA. This mixture was heat-shocked at 37°C for 30 seconds and left on ice for 90 minutes during which time it was shaken regularly. Four ml of LB medium was added to the tube which was incubated at 37°C for a further 90 minutes. The cells were then plated out, either 50, 100 or 200μl, onto LB agar plates containing an appropriate antibiotic.

Two controls were set up in addition. These were HB101 cells to which no plasmid DNA had been added, and HB101 to which a plasmid of known concentration had been added. After overnight incubation at 37°C, a single bacterial colony was picked from the probe DNA plate and inoculated into 10mls of LB medium with the appropriate antibiotic. The resulting overnight inoculant could then be used to make frozen stocks or to prepare large amounts of plasmid, as explained below.

### 2.8.3 Long Term Storage of Transformed *E. Coli*

Into sterile vials were placed 0.15mls sterile glycerol and 0.85mls fresh overnight culture. The two components were mixed and stored at -70°C. Viable cells could be recovered by simply scratching the surface of the frozen stock with a sterile metal loop.
2.9 Preparation of Closed Circular DNA using CsCl Gradients (Sambrook et al., 1989).

_E. Coli_ containing the desired plasmid or cosmid was streaked onto an LB agar plate containing antibiotic and incubated overnight. A single colony was used to inoculate 10mls of LB agar which was also incubated at 37°C overnight. 5mls of this culture was added to 250mls of LB with antibiotic and incubated at 37°C in an environmental shaker at 220 rpm. After 8 hours, 37.5mg of chloramphenicol was added to the culture which was incubated overnight. The cells were then pelleted at 4000g for 7 minutes and the pellet was resuspended in 10mls of Solution 1 (25mM Tris-HCl (pH 8.0) containing 50mM glucose, 10mM EDTA). 1ml of a freshly prepared solution of lysosyme (10mg/ml in 10mM Tris-HCl (pH 8.0)) was added and the mixture was left at room temperature for 10 minutes. 20mls of freshly prepared Solution 2 (0.2M NaOH, 1% SDS) was mixed in thoroughly and the solution was left on ice for 5 minutes. Finally, 15mls of Solution 3 was added and after mixing the contents were placed on ice for 30 minutes. (Solution 3 was made by adding 60mls 5M potassium acetate to 11.5mls glacial acetic acid and 28.5mls H₂O. The resulting solution is 3M with respect to potassium and 5M with respect to acetate.) The lysate was then centrifuged at 6000g for 10 minutes. The supernatant was poured through two layers of gauze into a 100ml cylinder and the volume noted. The supernatant was transferred to a clean centrifuge pot to which 0.6 volumes of isopropanol was added. This was mixed and stored at room temperature for 5 minutes. The nucleic acids were pelleted by centrifugation at 6000g for 10 minutes at room temperature. The supernatant was poured off carefully and the pellets were recentrifuged briefly at 400g. Any remaining supernatant was removed carefully with a pasteur pipette and the pellet was resuspended in 2.5mls of TE (pH 8.0). 500μl ethidium bromide from a 10mg/ml stock was added and the volume was made up to 5.0mls with TE (pH 8.0). Exactly 5.43g CsCl was added to the sample which was shaken until the CsCl dissolved. Any
protein-ethidium bromide complexes were then spun out of the mixture by cen­
trifugation at 12000g for 10 minutes. The clear supernatant was transferred to a 
Quick-seal centrifuge tube (Beckman). The remainder of the tube was filled 
with paraffin oil which had been warmed in a 37°C waterbath to facilitate pipet­
ting. Once the tube was full and any bubbles had been expelled, the tube was heat-sealed. The tube was centrifuged at 150,000g for 40 hours in a Beckman 
70.1Ti rotor at 20°C. After centrifugation, it was possible to see one or two pink bands when the tube was held against a white background in visible light. If two bands were present then only the lower band containing closed circular DNA was collected. The upper band, consisting of linear DNA, was collected first and discarded if it was not possible to collect the lower band without con­
tamination. In order to collect the bands, the tube was first secured in a clamp. 
The top of the tube was pierced with two 21 gauge needles to allow air to enter. 
A 21 gauge needle, attached to a 5ml syringe was inserted just below the band to be collected. The needle was turned bevel side up and the band was drawn into the syringe. The DNA was transferred to a 5ml tube and the ethidium bromide was removed by the addition of an equal volume of NaCl-saturated isopropanol. The tube was shaken and two phases formed upon settling. The ethidium bromide was in the upper, alcohol phase and was subsequently removed with a pipette. This process was repeated until no difference could be seen between the two phases.

In order to remove excess salt, which could subsequently inhibit enzyme activity, the DNA was dialysed overnight against 3 litres of TE (pH 8.0) which was changed twice during the dialysis. The DNA was transferred into a clean Corex tube and precipitated by the addition of one tenth volume of 3M sodium acetate and two (final) volumes of ethanol. The tube was left at -20°C overnight and the DNA pellet was recovered by centrifugation at 8000g for 10 minutes. The supernatant was poured off and the pellet was washed briefly in 80% ethanol. The DNA was re-pelleted, the supernatant poured off, and the tube was left inverted to drain. The pellet was dried off in a rotary evaporator (Savant)
and the pellet was resuspended in 250 - 500μl of TE (pH 8.0). The concentration of plasmid DNA was then determined by spectrophotometry as described in Section 2.11.

2.10 Preparation of DNA from Eukaryotic Cell Lines

To prepare DNA from blood or cell-lines for routine electrophoresis, the following method was used. Adherent cells were trypsinised to loosen them from tissue culture flasks. RPMI containing 10% FCS was then added to stop trypsin activity and prevent the cells from clumping. The cells were pelleted at 1000g for 5 minutes and washed once in PBS "A". The cell pellet was resuspended in 1-5mls of lysis buffer (10 mM Tris-HCl (pH 8.0), 10mM EDTA, 10mM NaCl, 0.5% SDS, 100μg/ml proteinase K) and incubated for 3-15 hours at 50 °C. The DNA was extracted with phenol once and phenol/chloroform twice. 20μl RNAse A was added (stock = 5mg/ml) and the DNA was incubated at 37°C for 30 minutes. The DNA was extracted with phenol/chloroform once and then chloroform at least once, following which the DNA was precipitated with 70% ethanol. The DNA was pelleted at 8000g for 10 minutes, washed in 80% ethanol and dissolved in TE (pH 8.0). The DNA concentration was determined as described in Section 2.11.

2.11 Spectrophotometric Determination of DNA Concentration

A 1/50 dilution of DNA was read on a spectrophotometer at 260nm and 280nm. The first reading was used to determine the DNA concentration: an optical density of 1 absorbance unit corresponds to approximately 50μg/ml for double-stranded DNA in a 1cm path cell. The ratio of the two readings ($A_{260}/A_{280}$) gave an indication of protein contamination. Pure preparations had a ratio of 1.8 whereas less pure preparations gave a significantly lower ratio, indicating that the $A_{260}$ reading may also be inaccurate.
2.12 Separation of Plasmid Insert and Vector DNA

There were several instances when it was preferable or necessary to prepare a sample of insert DNA free of plasmid vector sequence. This happened when the insert DNA was going to be radioactively labelled and the ratio of the size of the insert to the size of the vector DNA was 1 or less. Without removal of vector DNA, a large proportion of the radioactive DNA would be vector and not the sequence of interest.

To prepare insert DNA, 20 - 50µg of plasmid was digested with the restriction enzyme(s) corresponding to the insert site used in the construction of the plasmid. The digests were performed using the buffers recommended by the suppliers (New England Biolabs).Digests were carried out for 4 to 5 hours at the appropriate temperature using a 2 to 4 fold excess of restriction enzyme. If two enzymes were used the NaCl requirements of the enzymes were taken into consideration. If these were different, the DNA was first digested with the enzyme requiring a lower NaCl concentration. Additional NaCl was added and the second digest was performed. The total volume of enzyme added was not more than one tenth of the final digest volume in order to minimise non-specific endonuclease activity enhanced by glycerol in the enzyme storage buffer. The completeness of the digestion was then verified by running 1µg of digested DNA with size standards on a 0.8-1.0% agarose gel containing 0.5µg/ml ethidium bromide, using 1 x TAE buffer. The digest products were visualised under UV light. The size standards used were lambda DNA digested with the enzyme Hind III. If no undigested DNA was visible then the remaining digested DNA was separated by electrophoresis on a 0.8-1.2% low melting point agarose (SeaPlaque, FMC) gel using 1 x TAE buffer and ethidium bromide as described above. The DNA was then electrophoresed using a voltage gradient of not more than 3 V/cm. After electrophoresis, the gel was placed upon plastic film on a UV lightbox and the desired band was excised from the gel using a scalpel. This agarose block was trimmed to remove extraneous agarose and placed in a pre-weighed 1.5ml microtube. If the DNA was going to be used as a probe, three
volumes of water were added to the tube which was placed in a boiling water-bath for 10 minutes. The DNA was then stored at -20°C. To label this DNA with [³²P], the agarose sludge was defrosted and placed in a boiling waterbath for 5 minutes. 30µl was removed into a fresh tube and used in the labelling reaction which is described below. Alternatively, the DNA was separated from the agarose using the following method.

2.13 Glass Powder Isolation of DNA from Agarose Gels

A glass powder suspension was prepared as follows. 250mls of glass powder was resuspended in water in a 500ml beaker and allowed to settle for 1 hour. The supernatant was removed and centrifuged at 10000g. The supernatant was discarded and the pellet was resuspended in 100 - 200mls water. Nitric acid was added to a final volume of 50% and the mixture was brought to the boil. The mixture was centrifuged and washed four times with water. The resulting glass was stored as a 50% slurry in water.

A second solution was prepared by mixing 90.8g NaI with 1.5g Na₂SO₃ in 100ml water. The saturated solution was passed through a Whatman No.1 filter after which it was stored in the dark at 4°C.

A preweighed band of agarose containing DNA was cut with a scalpel into small pieces (1mm³). This was mixed in a 1.5ml microtube(s) with 2.5 volumes of the saturated NaI solution. The tubes were incubated at 45 - 55°C with occasional mixing for 5 minutes or until the agarose had totally dissolved. The glass slurry was meanwhile vortexed thoroughly. This was then added to the agarose solution. 5µl was added to solutions containing 5µg of DNA or less and a further 1µl was added for every additional 0.5µg of DNA above 5µg. The mixture was vortexed and placed on ice for 10 minutes. The tube was then centrifuged for 15 seconds in a microfuge (Microcentaur, MSE) at 6000g. The supernatant was removed with a pipette and the pellet was resuspended in 500 - 700µls of the cold ethanol wash solution (50% ethanol, 10 mM Tris-HCl pH7.5,
0.1 M NaCl, stored at -20°C). The suspension was vortexed as before and the supernatant was discarded. This wash step was repeated twice. After the supernatant had been removed from the final wash, the tube was re-centrifuged briefly and any remaining liquid was carefully removed with a fine pipette tip. The DNA was then eluted from the glass powder into a small volume of TE (pH 8.0). The DNA pellet was resuspended in half the desired final volume of TE and incubated at 45 - 55°C for 3 minutes. The suspension was then centrifuged for 40 seconds to form a hard pellet and the TE containing DNA was removed with a fine pipette tip into a fresh tube. This elution step was repeated with an equal volume of fresh TE and the two eluants were combined. The DNA yield was usually greater than 70 - 80%. The yield was frequently too small to be quantitated using absorbance spectroscopy and so was crudely assessed using the following method.

2.14 Quantitation of DNA using Ethidium Bromide Fluorescence

This method utilises the UV-induced fluorescence that ethidium bromide molecules emit when they are intercalated into DNA. The fluorescence of the sample is proportional to the mass of DNA present and so DNA samples of unknown concentration could be compared to a series of standards.

A piece of plastic film was smoothed onto a UV transilluminator and several 10μl aliquots of ethidium bromide (0.5μg/ml) were pipetted onto it. 1μl of a series of solutions containing lambda DNA at 5, 10, 30, 50 and 100μg/ml were mixed by pipetting into the ethidium bromide drops. 1μl of the sample DNA was mixed into a drop in the same manner. The drops were then illuminated and the fluorescence of the sample DNA was compared to the standards.

2.15.1 [³²P]-Labelling of DNA using the Random Priming Method

This protocol was taken from Feinburg and Vogelstein, 1983. 50ng of DNA in 30μl were boiled for 5 minutes, placed on ice for 2 minutes then centrifuged briefly. 2μl BSA (10mg/ml), 10μl OLB solution (see Section 2.18)
and 1 unit of Klenow was added and mixed by pipetting. 30μCi of [α-32P]dCTP was added immediately and mixed by pipetting. The reaction was left at room temperature for a minimum of 3 hours and more usually, overnight. The labelling mixture was then passed down a Sephadex-G50 column to remove any unincorporated deoxynucleotides, as described below.

2.15.2 Removal of Unincorporated [α-32P]dCTP using Sephadex-G50 Spin Column Chromatography

A 1ml syringe was plugged with siliconised glass wool and filled with Sephadex G50 (see Section 2.18). The syringe was placed in a narrow 15ml tube so that the grip ends of the syringe were supported on the rim of the tube. The tube was then centrifuged for 3 minutes at 300g, during which time the column packed down to between 0.9 - 1ml. 250μl of 2 x SSC was loaded onto the column which was re-centrifuged under the same conditions. The syringe was then transferred to a fresh 15ml tube at the bottom of which a 1.5ml microtube had been placed, devoid of cap. The 50μl labelling reaction was diluted with 200μl of 2 x SSC and loaded onto the column. Again, the column was centrifuged under the same conditions as before. The column, containing unincorporated deoxynucleotides was discarded and the 1.5ml microtube was recovered using a pair of forceps. The activity of the recovered DNA was determined by counting a 4μl aliquot in a bench-top β monitor.

2.16.1 Preparation of High Concentration Sonicated Human DNA for Preannealing Probes containing Repetitive Sequences

The source of DNA for this reagent was fresh placenta which was sliced into 2cm cubes and stored in 50ml tubes at -70°C. The DNA was prepared as follows.

A 50ml tube of tissue was allowed to thaw. Using forceps and sharp scissors, the tissue was snipped in small fragments into a 200ml bottle. 5mls of
10% SDS were added to the tube and the volume was made up to 100mls with STE (0.075M NaCl, 0.024M EDTA, pH 8.0). 100mg of proteinase K was added. The bottle was incubated for 40 hours in a 50°C waterbath, after which time no solid tissue was detectable in the bottle. The contents were then extracted twice with phenol, twice with phenol and chloroform and once with chloroform. The aqueous phase was transferred to a clean conical flask. The DNA was precipitated by the addition of 1/10 volume of 3M sodium acetate and 2 volumes of cold ethanol. The DNA was spooled out onto a glass rod and allowed to air dry. The DNA was then redissolved in 10mls of TE (pH 8.0) and vortexed thoroughly to aid dispersal of the pellet. A further 10mls of TE was added and the DNA was placed on a rotating platform at 4°C overnight. Once dissolved, the DNA concentration was determined using a spectrophotometer and the absolute amount of DNA present was determined. An average yield was 25mg per aliquot of placenta. The DNA was then sonicated for 6 minutes until it ceased to be viscous and air-bubbles rose swiftly through the solution. 1μg was loaded onto a 1% agarose gel with lambda/Hind III size standards and the average size of the DNA fragments was determined. If it was greater than 0.5 - 1 kb, the sonication step was repeated. The DNA was divided into two 30ml tubes and precipitated with 3M sodium acetate and ethanol as described above. The tubes were left at -20°C overnight. The tubes were then centrifuged in a cold centrifuge at 8000g for 20 minutes. The supernatant was discarded and the tubes were left to dry on the bench for 5 minutes, followed by 15 minutes in a rotary evaporator. Finally, the pellets were resuspended in TE (pH 8.0) to a final concentration of 10mg/ml.

2.16.2 Preannealing Probes Containing Repetitive Sequences

Some DNA markers contained both single-copy and repetitive sequences and it was necessary when probing filters to preanneal the repetitive sequences prior to hybridisation in order to avoid very high background "noise" in the DNA
tracks. The following method was used: The desired radiolabelled DNA marker was passed through a Sephadex-G50 column as described above and the activity was determined. The volume required for $10^7$ dpm was calculated. Into a 1.5ml microtube the following reagents were placed at the ratio to labelled probe indicated: 20 x SSC (1.2X), human sonicated DNA at 10mg/ml (1.2X), TE pH 8.0 (1.4X). The labelled probe was added to the tube last and the mixture was placed in a boiling water-bath for 10 minutes. Following this the tube was quickly transferred to a pre-heated dry-block (Dri-Block, Techne) at 65°C for 20 - 60 minutes; this time was determined by trial and error and varied between probes. After this time the entire mixture could be added to a pre-hybridised filter as described previously.

2.17 Isolation of DXS178 Positive Cosmids

2.16.1 Library Screening

The cosmid library ICRFcl00 was obtained from Dean Nizetic (ICRF, London). The library was constructed from $Mbo$ I partial digests of DNA from flow-sorted X chromosomes and ligated into the $BamH$ I cloning site of the cosmid vector Lorist 4. Lorist 4 is a derivative of Lorist B (Cross and Little, 1986). Duplicate sets of nylon filters representing over 8,000 colonies were screened with radiolabelled "212A". The isolation of the fragment "212A" has been described in Section 5 of this thesis. The filters were firstly incubated for 8 hours at 42°C in 20mls each of prehybridisation solution (50% formamide, 4xSSC, 50mM sodium pyrophosphate, pH 7.0, 10% dextran sulphate, 1% SDS, 50 μg/ml denatured salmon sperm DNA). The prehybridisation solution was then removed and to it was added 1x10⁶dpm/ml [$^{32}$P]-labelled "212A" which had been passed through a spin column and preannealed for 90 minutes with sonicated human DNA. In addition, 1x10⁶dpm/ml of $^{35}$S-labelled vector DNA (Lorist X, which differs from Lorist 4 in only a few bp of DNA) was added to the hybridisation solution. The filters were hybridised overnight and then washed to a final stringency of 1xSSC, 0.1% SDS at 65°C. The filters were
autoradiographed as usual except that a thin sheet of plastic was placed between
the filters and the film to allow only a weak signal from the $^{35}$S to penetrate.
This produced a weak background signal from each colony and allowed the grid
coordinates of $[^{32}\text{P}]-212\text{A}$ positive colonies to be determined. Seventeen
colonies gave signals which were strong compared to the background signal (16
on duplicate filters). The coordinates were sent to ICRF, London where the
master colonies were stored.

2.17.2 Primary Screening

The clones were received from ICRF, London as transformed bacterial stabs.
The host bacteria was E. Coli strain ED8767 (Murray et al., 1977). The
colonies were originally grown in 96-well microtitre plates and the automated
method by which they were spotted onto the nylon filters was such that colonies
from adjacent wells of the same microtitre plate were represented by colonies
spaced 5 apart on the nylon filters. Several of the 212A positive clones were
spaced in an identical fashion from each other, indicating that the clones
originated from adjacent wells and that there may have been cross-contamination
between wells. The primary step in screening the "positive" colonies was
therefore to establish whether colonies contained mixed populations of clones
and whether they were truly positive. The stabs were streaked out onto LB agar
with kanamycin (3µg/ml) and incubated at 37°C overnight. Multiple colonies
were picked from these plates (4 from plates with suspected mixed populations;
two from the others) into 10mls LB with kanamycin and cultured overnight.
5mls of the bacterial suspensions were pelleted and closed circular DNA was
prepared using QIAGEN columns (Hybaid, UK). The DNA was digested with
EcoRI or BamHI, separated on a 0.8% agarose gel and blotted. The filters were
hybridised with 212A. Duplicate and mixed clones were detected using the
restriction and hybridisation patterns obtained.
2.17.3 Analysis of DXS178-Positive Cosmids using PFGE

The PFGE separation conditions were as follows: DNA was separated using the CHEF adaption on a 1% agarose gel at 190V using a 2 second switch time. The gels were run for 15 hours and the buffer temperature was 13°C. The markers used were bacteriophage lambda concatamers and lambda digested with the enzymes Sal I, Apa I or Hind III.

2.18 Reagents

All reagents were made using water purified by reverse osmosis.

10xTris-borate (10xTBE)
108g Tris base, 55g boric acid, 20ml 0.5M EDTA (pH8.0), water to 1 litre.
Autoclave at 121°C for 15 minutes.

20XSSC
175.3g sodium chloride, 88.2g sodium citrate, water to 1 litre.

Tris-EDTA (TE) pH8.0
To 1 litre water which had been autoclaved at 121°C for 15 minutes, the following was added: 10mls 1M Tris-HCl (pH8.0), 2mls 0.5M EDTA (pH8.0).

Denhardt's Solution (100X)
20g Ficoll 400 (Pharmacia Biosystems), 20g polyvinylpyrrol-idone, 20g bovine serum albumin (fraction V), water to 1 litre. This was filtered and stored at -20°C.

LB (Luria-Bertani) Medium
10g Bacto-tryptone (Difco), 5g Bacto-yeast extract (Difco), 10g NaCl, water to 1 litre. The mixture was autoclaved for 20 minutes at 121°C.
YEPD

10g Yeast Extract (Difco), 20g peptone, 20g glucose, water to 1 litre. The mixture was autoclaved as for LB.

LB and YEPD Agar Plates

To the appropriate media, 15g/l Bacto-agar was added before autoclaving.

Antibiotics

Ampicillin

The stock solution was 25mg/ml of the sodium salt in water, filter-sterilised and stored at -20°C.

Kanamycin

The stock solution was 25mg/ml in water, filter-sterilised and stored at -20°C.

Tetracycline

The stock solution was 12.5mg/ml tetracycline hydrochloride in ethanol/water (50% v/v), stored in the dark at -20°C.

The working concentrations for the above antibiotics were 1/1000 of the stock solutions.

Preparation of Sephadex G-50 for Spin Column Chromatography

10g of Sephadex G50 was added to 100mls sterile 2xSSC in a bottle. This was incubated at 65°C for three hours during which time the supernatant was replaced with new 2xSSC. The prepared Sephadex was stored at 4°C.

Preparation of OLB Solution for Radiolabelling using the Random Priming Technique

OLB was made from the following components:

Solution O: 1.25M Tris-HCl, 0.125M MgCl₂ at pH 8.0 (stored at 4°C).
Solution A: 1ml Solution O, 18μl 2-mercaptoethanol, 5μl each of dATP, dTTP and dDTP which had been previously dissolved in TE at 0.1M concentrations (stored at -20°C).

Solution B: 2M Hepes pH6.6 (stored at 4°C).

Solution C: Hexanucleotides (Pharmacia Biosystems) suspended in TE at 90 OD units/ml (stored at -20°C).

OLB was made from mixing solutions A:B:C at a ratio of 10:25:15 and stored at -20°C.
3. RESULTS: MAPPING PROBES BY DELETION AND LINKAGE ANALYSIS

3.1. Localisation of Probes within Xq21.3-Xq22 by Deletion Mapping

In order to increase the likelihood of physically linking the polymorphic probes that had previously been used in XLA family studies (see section 1.4.2), a number of additional probes were used. These included DXS114 (757) and DXS83 (716) which had been previously localised to Xq21.3 - Xq22 (Mandel et al., 1989). In order to refine the localisations of these markers, they were hybridised to Southern blots containing enzyme-digested DNA from a normal male and female and also from a lymphoblastoid cell line from TCD patient MBU whose deletion has been described in Section 1.4.1.3. The markers DXS178 and DXS87 were also hybridised to this panel in an attempt to refine their previous localisations (Mandel et al., 1989). Details and references for all the probes used in this study are given in Appendix A1. Additional mapping studies were carried out with probe DXS114 on DNA from the lymphoblastoid cell lines DM and NP (see Section 1.4.1.3). The results are shown in Figure 3.1 and summarised in Table 3.1.

Probes DXS83, DXS87 and DXS178 gave signals in the DNA from normal male and female DNA and also in the MBU track. As the MBU deletion encompasses Xq21.2-q21.33, the locations of DXS83, DXS87 and DXS178 can be refined to Xq22.

Probe DXS114 did not hybridise to DNA from MBU, indicating that this marker was localised within the deletion of patient MBU. Further mapping showed that DXS114 was also absent from NP but present in the DM track, indicating that it maps to the proximal part of Xq21.3.
c. Positive control for Figure 3.1b. NP is positive for DXS118, and gives a signal comparative in strength to the DM track. This contrasts with Figure 3.1b, where the NP track shows non-specific hybridisation only compared with the positive signal in the DM track.
Figure 3.1
Localisation of DNA Markers using Deletion Mapping Panels

F = normal female DNA; M = normal male DNA. Cell lines NP, MBU and DM have been described in Section 1.4.4. Cell lines AnLy and RAG have been described in Section 2.1. DNA was digested with a) EcoRI or b) Pst I and separated on an 0.8% agarose gel. Southern blotting and hybridisation was carried out as described in Materials and Methods. Size markers (kb) are bacteriophage λ digested with Hind III.
### Table 3.1
Presence of Probes in Cell Lines from TCD Patients

<table>
<thead>
<tr>
<th>DNA Probe</th>
<th>HGM 10 Location</th>
<th>46XY</th>
<th>46XX</th>
<th>MBU</th>
<th>NP</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS114</td>
<td>Xq21.3-q22</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>DXS83</td>
<td>Xq21.3-q22</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>DXS87</td>
<td>Xq21.3-q22</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>DXS178</td>
<td>Xq21.3-q22</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>nd</td>
<td>nd</td>
</tr>
</tbody>
</table>
3.2. Linkage Studies with DXS101

3.2.1 Analysis of Recombinant Chromosomes using DXS101

DXS101 has been previously localised to Xq22 by deletion analysis (Cremers, 1988), but has not been used in any linkage studies published to date. Previous analysis of families with X-linked syndromes (Goodship, 1989b) has identified at least three chromosomes with recombinations that have occurred between DXS3 (in Xq2.13) and probes mapping to Xq22. The families in which two of these cross-overs occurred were found to be informative for DXS101 in this study and so were analysed with a view to ordering DXS101 relative to other probes mapping to Xq22.

In family HIS (Figure 3.2) a recombination has occurred between DXS3 and DXS17 in individual II.4. The DXS101 alleles are in phase with those of DXS17. The simplest explanation for this is that DXS101 maps distal to DXS3, but it is not possible to position DXS101 relative to DXS17 on the basis of this data alone. The second family, MAN, has three individuals affected with XLA. In individual II.1 a recombination has occurred between the disease locus and DXS3. Probe DXS178 has been previously used to investigate this family and this probe was found to segregate with the XLA locus (Goodship, 1989b). In the present study, DXS101 was also found to segregate with DXS178 and the disease locus. Once more, DXS101 is positioned distal to DXS3 but it is not possible to position DXS101 relative to the XLA locus or to DXS178.

3.2.2 Linkage Analysis of XLA Families using DXS101

DXS101 was used to perform linkage analysis in several XLA families. Out of 29 unrelated chromosomes tested, 11 carried the upper 7.7kb allele and 18 carried the lower 7.5kb allele. This gives allele frequencies of 0.38 and 0.62, respectively, which compare well with the published frequencies of 0.35 and 0.65, respectively (Mandel et al., 1989).
Figure 3.2
Linkage Analysis of Families with Recombinations in Xq21.3-q22

Family HIS

```
DXS3  +  
DXS17  +  
DXS101 -  
```

```
+  + 
II.1  +  II.2  -  II.3  + 
+  + 
```

Family MAN

```
DXS3  +  
DXS178 -  
DXS101 -  
```

```
+  +  +  +  
II.1  II.2  II.3 
```

Key:  
- Male  
- Female  
- Affected male  
- Obligate carrier  
 Indicates most probable site of recombination  
 + Indicates presence of polymorphic site  
 - Indicates absence of polymorphic site
Figure 3.3
Linkage Analysis of XLA Families using DXS101

Family MON

Family MAN

Family KIR

Family GER

Family PAT

For key see figure 3.2.
DXS101 was used to analyse five informative XLA families, including the previously mentioned MAN family (Figure 3.3). No recombinations were found between DXS101 and the XLA locus and a maximum lod score of 3.01 was obtained at $\theta = 0$ with 95% confidence limits of $0 < \theta < 0.2$.

3.3 Discussion

*Mapping Probes using Lymphoblastoid Cell Lines*

The reliability of the deletion mapping data in Section 3.1 is based upon the assumption that each TCD patient has a single, continuous deletion. As all of these cell lines have previously been used in combination to localise probes in proximal Xq, without any discrepancies to indicate that the cytogenetic deletions are caused by loss of more than one DNA fragment (Cremers *et al.*, 1988; 1989), this is a fair assumption. The chromosome bands to which the probes have been localised are illustrated in Figure 3.4.

DXS83, DXS87 and DXS178 are all present in MBU and therefore by combining this data with the previously published information that they map to Xq21.3 - Xq22, they must be located in Xq22. In previous studies, DXS87 has been shown to overlap DXS88 (pG3) by PFGE studies and by virtue of the fact that the two probes recognise the same polymorphism (Goodship, 1989b). DXS88 was originally assigned to Xq21.33 (Mandel *et al.*, 1989). As MBU is deleted for Xq21.33, the original localisation of this marker appears to be proximal to the most likely location for DXS88 in Xq22.

*Linkage studies with DXS101*

The marker DXS101, when used for linkage analysis in XLA families, produced a maximum Lod score of 3.01 at $\theta = 0$. A positive Lod score was anticipated as DXS101 has already been assigned to Xq22 (Cremers *et al.*, 1988). The tracking of DXS101 in patients with recombinations and in XLA families did not enable the position of DXS101 relative to other Xq22 probes to be elucidated. The published allele frequencies of DXS101 (0.65 and 0.35)
Figure 3.4
Refinement of the Loci for DXS83, DXS87, DXS114 and DXS178 based on Deletion Mapping and Previous Assignments
indicate that 46% of females should be heterozygous for this marker and therefore this probe could be useful for genetic counselling in families which are not informative for probes DXS178, DXSS94, DXS17 or DXS3. These four probes have already been found to be separated from the XLA locus by a maximum genetic distance of 6.5 cM (see section 1.4.11). These four markers are used for prenatal testing in XLA families, so should continue to be used in preference to DXS101 which could be situated distal to DXS17 and therefore at a greater distance from the XLA locus. In addition, the close proximity of the polymorphic alleles of DXS101 on agarose gels due to their similarity in size (7.7kb and 7.5kb) means that for good separation the DNA should be electrophoresed for 2 days instead of the standard one day, adding to laboratory procedure time.
4. RESULTS: PHYSICAL MAPPING OF Xq22 BY PFGE AND ANALYSIS OF RADIATION HYBRIDS

4.1 Physical Mapping of Xq22 by Pulsed Field Gel Electrophoresis

In order to obtain a physical map of the XLA locus, long-range mapping studies were carried out using PFGE. The polymorphic markers DXS178, DXS94, and DXS17 that have been used for linkage studies in XLA families were employed in addition to seven probes which had been mapped, either previously or in this study, to Xq22. Details and references for the probes used are given in Appendix A1.

Initially, the restriction enzymes Sfi I, Mlu I, BssH II and Sal I were used to digest female leukocyte DNA which was then separated by PFGE. The gels were blotted and sequentially hybridised with the various Xq22 markers and the fragment sizes obtained are shown in Table 4.1.

DXS101, which has been used for linkage analysis in this study (Section 3.2) has not been included in Table 4.1 because it is known to recognise at least five discrete DNA species in Xq22 (Hofker et al., 1987). Preliminary PFGE mapping with this probe produced at least four Sfi I bands and four Sac II bands, indicating that these DNA species are dispersed over a minimum of 1Mb DNA and therefore likely to give ambiguous physical linkage results.

A significant gel to gel variation was found between the mobility of the yeast markers and genomic DNA tracks, making it difficult to compare band sizes between different gels. For example, DXS94 gave Sfi I bands of 650kb and 540kb on one gel (Figure 4.12) but of 680kb and 570kb on a second gel (Figure 4.13). In addition, DNA fragments which did not lie within the window of optimal resolution for particular gels could only be sized very inaccurately. This was especially true for fragments less than 1Mb which had been resolved on gels whose run conditions were designed to separate up to 3Mb. Therefore, whenever probes are indicated in a table as recognising similar sized fragments with a particular restriction enzyme, this will have been verified by hybridising
Table 4.1
Fragments Detected in Female Leukocyte DNA using PFGE

<table>
<thead>
<tr>
<th>DNA Locus</th>
<th>Sfi I</th>
<th>Bss H II</th>
<th>Mlu I</th>
<th>Sal I</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS178</td>
<td>130</td>
<td>470</td>
<td>130</td>
<td>several</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>330</td>
</tr>
<tr>
<td>DXS94</td>
<td>570</td>
<td>380</td>
<td>2700</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>650</td>
<td>560</td>
<td>670</td>
<td></td>
</tr>
<tr>
<td>DXS147</td>
<td>570</td>
<td>930</td>
<td>2700</td>
<td>nd</td>
</tr>
<tr>
<td></td>
<td>650</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS211</td>
<td>150</td>
<td>930</td>
<td>2700</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>230</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>650</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS17</td>
<td>215</td>
<td>930</td>
<td>2700</td>
<td>330</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>400</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>450</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>560</td>
</tr>
<tr>
<td>DXS87</td>
<td>400</td>
<td>830</td>
<td>2700</td>
<td>330</td>
</tr>
<tr>
<td></td>
<td>760</td>
<td>970</td>
<td>450</td>
<td></td>
</tr>
<tr>
<td>PLP</td>
<td>130</td>
<td>260</td>
<td>590</td>
<td>460</td>
</tr>
<tr>
<td></td>
<td>260</td>
<td></td>
<td>580</td>
<td></td>
</tr>
<tr>
<td>DXS54*</td>
<td>&lt;=90</td>
<td>125</td>
<td>1060</td>
<td>460</td>
</tr>
<tr>
<td></td>
<td>240</td>
<td>340</td>
<td>815</td>
<td>580</td>
</tr>
<tr>
<td></td>
<td>430</td>
<td>510</td>
<td>630</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>640</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS24</td>
<td>610</td>
<td>570</td>
<td>1060</td>
<td>580</td>
</tr>
<tr>
<td>DXS83</td>
<td>610</td>
<td>570</td>
<td>1060</td>
<td>520</td>
</tr>
</tbody>
</table>

Fragment sizes are in kb. Where more than one band is indicated, the larger fragments are partial digest products.

* DXS54 recognises an autosomal fragment in addition to an X-linked sequence and the fragments listed include all those detected.
both probes to the same filter or to sections of the same filter. In addition, the measurements will have been taken from a gel which included the fragment size in its optimal separation range.

The band sizes shown in Table 4.1 indicate that a number of recognise similar sized DNA fragments derived with particular restriction enzymes. Examining the \textit{Mlu} I and \textit{Sal} I fragments produces evidence of two major linkage groups: the first group is DXS17, DXS87, DXS94, DXS147, and DXS211 and the second group is PLP, DXS24, DXS54 and DXS83. In order to minimise the possibility of two probes being wrongly linked together by virtue of the fact that they recognised similar sized but separate bands with a particular restriction enzyme, these possible linkage groups were tested more rigorously using further digests and different sources of DNA. If two probes were found to both recognise at least two different bands, either with two or more restriction enzymes or with complete and partial digest products with a particular enzyme, this was taken as evidence of physical linkage.

4.1.1 Linkage of Probes DXS24, DXS54, DXS83 and PLP

Probes DXS54 and PLP both recognised \textit{Sal} I bands of 460kb and 580kb in female leukocytes, while DXS24 recognised a single 580kb band and DXS83 detected a 520kb band (Figure 4.1). This indicated that DXS54 and PLP were linked on the same 460kb \textit{Sal} I fragment, as they both recognised an additional 580kb partial product (neither of these probes contained an internal \textit{Sal} I site). The low probability that the two probes each recognised two different \textit{Sal} I bands of coincident size was higher for these two markers than for two single copy probes, because St3 (DXS54) recognises an additional autosomal fragment in addition to its X chromosome specific fragment. Also, PLP is a cDNA clone and therefore could possibly detect fragments which span a \textit{Sal} I site in an intron. In order to determine the relative intensities of the X-linked and autosomal fragments recognised by DXS54, this marker was hybridised to a Southern blot containing DNA from cells containing 1, 2 and 4 X chromosomes.
DNA from male (M), female (F), LCL-127 (4X), Rag (R) or AnLy (A) cells was digested with either EcoRI (tracks 1-7) or Hind III (tracks 8 - 14). In tracks 1-5 and very faintly in 7, a 2.1 kb band (X) can be detected. This is expected as DXS54 is a 2.1 kb Eco RI genomic fragment. In addition, larger bands (A) are visible in the male tracks of a similar intensity to the 2.1kb fragment. In the 2X tracks, the 2.1kb fragment is the same intensity as in the male tracks, indicating that the female tracks are less heavily loaded than the male. No larger bands are visible. In the 4X track, the 2.1 kb band is of 3-4 fold greater intensity than in the male tracks as estimated by eye, whereas the larger bands in the 4X track are of comparative intensity to those in the male tracks. This indicates that the signal from the two autosomes in the male DNA is approximately equivalent in intensity to the X-specific signal from a single X chromosome. Tracks 7-11 show the same DNA samples, this time digested with Hind III. The DXS54 plasmid has no internal site for Hind III (data not shown). The intensity of the largest band (X) increases in proportion to the number of X chromosomes present in each sample, although again the 2X tracks which are underloaded. In contrast, the intensity of the additional bands in tracks 8 -12 (A) reflects the loading of the autosomal DNA in each track. The larger of these two bands also appears to be visible in the AnLy track (14), however, a similar band is present in RAG (13), indicating that it is of rodent origin. The relative intensities of the X-linked and autosomal bands in the 2X DNA is such that the strongest autosomal signal is only half as strong as the X-specific signal. Size markers are λ/Hind III.
DNA from female leukocytes was digested with SalI and separated on a 1% gel using CHEF. Run conditions were 170 V with a switch time of 80 s. Size markers (kb) are YP148 chromosomes.
and also from the somatic cell hybrid AnLy which contains Xq as its only human component (see Figure 4.2A). The signal from a pair of autosomes was found to be comparative in intensity to that from a single X chromosome. In Figure 4.1, both SalI bands detected by DXS54 in female, 2X DNA were of similar intensity; therefore they were both likely to be X-linked. This substantiates the proposed linkage linkage between DXS54 and PLP.

PLP was found to recognise a 590kb Mlu I band in female leukocyte DNA. DXS54, DXS24 and DXS83 all hybridised to a 1050kb Mlu I fragment and DXS54 recognised additional bands of 820kb and 640kb (Table 4.1, Figure 4.2). In order to determine which of the bands detected with DXS54 were X-specific, these digests were repeated using the cell lines LCL-127 and AnLy. The results are shown in Table 4.2 and in Figures 4.2 and 4.3.

PLP gave a 590kb band with AnLy, the same size as the band detected in normal female leucocytes, whereas DXS54 recognised a separate band at about 200kb and DXS83 a band at around 850kb with AnLy (Figure 4.3). This may reflect the presence of an additional Mlu I site in this cell line, one which is normally methylated in leukocytes and which therefore gives rise to a larger 1060kb band recognised by both DXS54 and DXS83. Both PLP and DXS54 hybridised to DNA fragments of 500kb and 590kb with similar levels of intensity in the LCL-127 tracks (Figure 4.2), which supports the Sal I physical linkage data for these two probes. DXS24 and DXS83 both recognised Mlu I bands of 430, 590, 680 and 1400kb in LCL-127 and a band of 1060kb in female leukocytes (Figure 4.2). They also detected identically sized BssH II fragments of 570kb (Table 4.1; Figure 4.4) and Sfi I fragments of 610kb (data not shown). The fragments detected in the AnLy tracks in Figure 4.4 have migrated slightly further than the bands in the female leukocyte tracks. This may well be caused by a DNA loading effect, rather than different digestion sites within the AnLy DNA. In Figure 4.2, it can also be seen that all four markers apparently detected a similar sized diffuse band of approximately 1400kb in the LCL-127 tracks. This corresponds to a part of the gel where resolution is reduced, so that
DNA from 4X cell line LCL-127 (4X) and female leukocytes (F) was digested with Mlu I and separated on a 1% gel using CHEF. Run conditions were 170 V with a switch time of 80 secs for 26 h followed by 180 secs for 14 h. The buffer temperature was 14 C. Size markers (kb) are YP148 chromosomes. The four panels show autoradiographs of the same filter which was sequentially hybridised. The faint bands in the DXS54 panel are from DNA species of autosomal origin (see Figure 4.2A).
Table 4.2
Fragment sizes obtained using *Mlu I* on different DNA samples with PLP, DXS24, DXS54, and DXS83

<table>
<thead>
<tr>
<th>DNA Locus</th>
<th>46XX</th>
<th>LCL-127</th>
<th>AnLy (Xq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLP</td>
<td>590</td>
<td>390</td>
<td>590</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>590</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1350</td>
<td></td>
</tr>
<tr>
<td>DXS54</td>
<td>1060</td>
<td>500</td>
<td>210</td>
</tr>
<tr>
<td></td>
<td></td>
<td>590</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>730</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1400</td>
<td></td>
</tr>
<tr>
<td>DXS83</td>
<td>1060</td>
<td>430</td>
<td>850</td>
</tr>
<tr>
<td></td>
<td></td>
<td>590</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>690</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1350</td>
<td></td>
</tr>
<tr>
<td>DXS24</td>
<td>1060</td>
<td>430</td>
<td>nd</td>
</tr>
<tr>
<td></td>
<td>1450</td>
<td>590</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>690</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1350</td>
<td></td>
</tr>
</tbody>
</table>

Fragment sizes are in kb. nd = not done.
DNA from female leukocytes (F) or cell line AnLy (A) was digested with BssH II and separated on a 1% gel using CHEF. Run conditions were 170V with a switch time of 80 secs for 18h followed by 180 secs for 24 hours. The buffer temperature was 15°C. Size markers (kb) are YP148 chromosomes. The slightly faster migration of the bands in the AnLy tracks probably reflects the slightly lower DNA concentration in these blocks.
markers in the range 1200kb to 1500kb are compressed into approximately 8mm. This fragment is in the correct size range for a partial \textit{Mlu} I digest product if one adds the sizes of the next-largest fragments seen in the LCL-127 tracks with PLP/DXS54 (590 kb) and DXS83/DXS24 tracks (690 kb).

DXS54 detected separate sized \textit{BssH} II fragments (Figure 4.5) and \textit{Sfi} I fragments (Figure 4.6) to DXS24 and PLP.

DXS24 hybridised to a 580kb \textit{Sal} I fragment, however this is unlikely to be the same fragment as the partial product recognised by DXS54 and PLP because DXS24 did not detect a small 120kb fragment as would be expected if it was recognising a 580kb partial product.

Assuming that the linkage of DXS54 to DXS24 and DXS83 on an \textit{Mlu} I fragment is correct, the four probes can be ordered as (DXS83,DXS24) - DXS54 - PLP. The relative order of DXS83 and DXS24 has not been determined in this study, however it should be possible to determine with \textit{Sal} I partial digests as these probes hybridise to dissimilar bands of 520kb and 580kb, respectively (Figure 4.3).

The mapping data for these four probes is summarised in Figure 4.7. The maximum distance over which these probes are positioned is as follows. The three probes DXS24, DXS54 and DXS83 are contained on a single 1060kb \textit{Mlu} I fragment. DXS54 and PLP recognise identical 500kb \textit{Mlu} I bands on LCL-127. These four probes are therefore contained on DNA spanning approximately 1600kb.

4.1.2 Linkage of DXS17, DXS94, DXS147 and DXS211

The five markers DXS17, DXS87, DXS94, DXS147 and DXS211 were all found to hybridise to a 2700kb \textit{Mlu} I fragment (Figures 4.8 and 4.9). DXS147, DXS17 and DXS211 all detected a 930kb \textit{BssH} II band which was reduced to 720kb upon double digestion with \textit{Mlu} I (Figures 4.10 and 4.11). DXS94 detected a separate \textit{BssH} II band of 380kb with partial products of 560kb and 670kb. DXS94 and DXS147 both hybridised to a 570kb \textit{Sfi} I fragment and
Figure 4.5
PLP, DXS24 and DXS54 Detect Separate BssH II Fragments

DNA from 4X cell line LCL-127 (4X) and from female leukocytes (F) was digested with BssH II and separated on a 1% gel using PFGE. Run conditions were 170 V with a switch time of 80 secs for 16 h followed by 180 secs for 24 hours. The buffer temperature was 14°C. Size markers are YP148 chromosomes.
PLP, DXS24 and DXS54 are Located on Separate Sfi I Fragments

DNA from 4X cell line LCL-127 (4X) and from female leukocytes (F) was digested with Sfi I and separated on a 1% gel using CHEF. Run conditions were as described in Figure 4.5. Size markers (kb) are YP148 chromosomes.
Figure 4.7
Summary of Physical Linkage Data for PLP, DXS24, DXS54 and DXS83

This diagram was compiled using data obtained from female leukocyte DNA and AnLy DNA. Large brackets [ ] indicate that the orientation of particular markers or probes have not been determined. Small brackets ( ) indicate a restriction site that is only present in AnLy DNA.
DNA from female leukocytes was digested with *Mlu* I and separated on a 1% gel using CHEF. Run conditions were 40V with a switch time of 1900 secs for 240 h. The buffer temperature was 8 C. Size markers (Mb) are *H. wingei* chromosomes. Key: M = *Mlu* I. The four panels in this figure represent twelve adjacent tracks from the same gel.
Figure 4.9
Detection of *Mlu I* Fragments by DXS94, DXS178 and DXS211

DNA from female leukocytes was digested with *Mlu I* and separated on a 1% gel using CHEF. Run conditions were 40V with a switch time of 2000 secs for 180 h. The buffer temperature was 14°C. Size markers (Mb) are *H. wingei* chromosomes. Key: M = *Mlu I*. 
DNA from female leukocytes was digested with Bss H II and separated on a 1% agarose gel for 40 h. The switch time was 120 secs, the voltage was 120V and the buffer temperature was 14°C. Size markers (kb) are YP148 chromosomes. Key: B = BssH II.
Figure 4.11
Detection of $Bss\, H\, II$ and $Bss\, H\, II/ Mlu\, I$ Bands with DXS94, DXS147 and DXS211

DNA from female leukocytes was digested with $Bss\, H\, II$ or $Bss\, H\, II + Mlu\, I$ and separated using conditions described in Figure 4.10. Key: $B = Bss\, H\, II$, $M = Mlu\, I$. 
to a fainter 650kb partial digest product (Figure 4.12). This suggests that DXS147 is flanked on one side by DXS94 and on the other by DXS211 and DXS17. DXS211 recognises a 650kb \textit{Sfi} I partial product in addition to smaller bands (Figure 4.13) whereas DXS17 does not (Figure 4.12). This is the same size as the larger \textit{Sfi} I fragment recognised by DXS94 and DXS147.

By performing double digests with the enzymes \textit{BssH} II and either \textit{Sfi} I (Figure 4.12) or \textit{Mlu} I (Figure 4.11), it was possible to map the DNA fragments recognised by DXS17, DXS94, DXS147 and DXS211 relative to each other and a physical map featuring these probes is shown in Figure 4.14.

This data implies that probes DXS17, DXS94, DXS147 and DXS211 are physically linked on the same 2.7Mb \textit{Mlu} I fragment with the order DXS94-DXS147-DXS211-DXS17. From genetic linkage data (Section 1.4.2), DXS94 is centromeric to DXS17, therefore this fragment can also be orientated with respect to the centromere.

DXS87 does not share similar \textit{Sfi} I or \textit{BssH} II fragments with any of these four probes (Table 4.1). The two \textit{Sal} I fragments detected by DXS87 are the same size as two of the fragments recognised by DXS17 (data not shown). However, as DXS87 does not recognise the largest 560 kb band that is detected by DXS17, it is assumed that the similarity in these band sizes is coincidental. This data suggests two possibilities for the location of this probe, which are:

1) DXS87 is on the same 2.7Mb \textit{Mlu} I fragment as these probes and must lie centromeric to DXS94. DXS87 cannot be telomeric as the the 930kb \textit{BssH} II fragment recognised by the other probes overlaps the telomeric end of the \textit{Mlu} I fragment.

2) DXS87 is located on a different 2.7Mb \textit{Mlu} I fragment to that which is recognised by DXS17, DXS94, DXS211 and DXS147.

4.2 Orientation of Xq22 Probes using Radiation Hybrids

Using PFGE it was possible to determine that eight of the DNA markers used were physically linked into two blocks. It was not possible to link these
Figure 4.12
Detection of Sfi I and Sfi I / Bss H II Fragments with DXS17, DXS87, DXS94 and DXS147

DNA from female leukocytes was digested with Sfi I and/or BssH II and separated on 1% gel using CHEF, with a switch time of 70 secs at 170V for 40 h. The buffer temperature was 11°C. Size markers are YP148 chromosomes. Key: S = Sfi I, B = BssH II.
DNA from female leukocytes was digested with \textit{Sfi} I and separated using CHEF for 70 h, using a switch time of 70 secs at 120V. The buffer temperature was 14°C. Size markers (kb) are YP148 chromosomes.
Figure 4.14
A Physical Map around the Loci DXS17, DXS94, DXS147 and DXS211

Scale (Mb)

Key: B = Bss H II, S = Sfi I, M = Mlu I.
* The relative positions of these two sites have not been determined.
groups together, however, or to determine their relative orientation with respect to each other. Also, it was only possible to orientate one group, i.e. that which contained DXS94, with respect to the centromere by using information from linkage studies.

The use of radiation hybrids to order probes has been discussed in Section 1.6.3. To compliment the PFGE analysis of Xq22, X chromosome hybrids were also analysed with probes from this region, with the intention of orientating the groups of linked probes. The IFGT hybrids used were kindly provided by Peter Goodfellow (ICRF, London) and have previously been described (Benham et al., 1989; Section 1.6.3). Eight hybrids were chosen for analysis on the basis that on a primary screening panel provided by Peter Goodfellow, they were all positive for DXS17: of these, seven were DXYS5 negative and five were DXS11 negative (see Appendix A2). DXYS5 has been localised to Xq21 and DXS11 to Xq24-q26 therefore this choice of hybrids selected for clones which contained human DNA with a fragmentation point somewhere in Xq21.3-q22 and/or in the region Xq22-Xq26.

The hybrids were grown and DNA was prepared as described in Section 2.1. DNA was digested, electrophoresed, Southern blotted and then hybridised with various Xq22 probes. The restriction enzymes used were chosen to correspond to the cloning sites of the various Xq22 probes. This was important in helping to determine which bands were human-specific by size, as several probes cross-hybridised to DNA species in the hamster parent lines. The controls included on all filters were C12D, which is the parent line containing a single human X chromosome and Wg3-H, which is the rodent parent of C12D. Unfortunately it was not possible to obtain cell line A23, which was the hamster fusion partner for C12D and which would have been more suitable as a negative control.

The eight hybrids were scored for the presence or absence of Xq22 probes and the results can be seen in Table 4.3. DXS17 was rescored in this study to rule out the possibility that the hybrids had lost human chromosomal material since their primary screening.
Table 4.3
Orientation of Xq22 Probes using Radiation Hybrids

<table>
<thead>
<tr>
<th>Clone No.</th>
<th>DXS178</th>
<th>DXS94</th>
<th>DXS147</th>
<th>DXS211</th>
<th>DXS17</th>
<th>PLP</th>
<th>DXS83</th>
<th>DXS87</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>52</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+wik</td>
</tr>
<tr>
<td>92</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>129</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>143</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>203</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Key:  
+ = positive hybridisation of DNA marker to clone DNA  
- = no hybridisation of DNA marker to clone DNA  
+wik = weak hybridisation
The method by which the data was analysed was to arrange the Xq22 loci so that the score results from the clones were compatible with each clone having only one human insert containing Xq22 probes. The clones have been arranged in this way in Table 4.3 taking into account all the scores except those for DXS87. By simply looking at the positive scores for the remaining probes, it can be seen that this probe order is consistent with the presence of a single Xq22 fragment in all eight clones. The loci DXS147, DXS211 and DXS17 can be interchanged without affecting this, indicating that these probes are very close together, which has been demonstrated by physical linkage in section 4.3. A prior constraint which has been placed upon the probe order in Table 4.3 is that DXS178 is centromeric to DXS94, in agreement with the linkage data described in Section 1.4.1.1. Of all the probes used in this study DXS17, DXS147 and DXS211 were found to be present in all eight hybrids, suggesting that these markers lie close to each other. PLP and DXS83 cannot be centromeric to DXS178, as this would necessitate the presence of more than one fragment in clones 52, 92 and 203. By ignoring DXS94, PLP may be placed between DXS178 and DXS17, but DXS83 cannot because this is contradicted by clone 129. If DXS94 is placed centromeric to DXS17, as is inferred from linkage data, then clone 52 contradicts the positioning of both DXS83 and PLP between DXS178 and DXS17, indicating that they lie distal to DXS17.

The orientation of DXS83 and PLP with respect to the centromere must be with PLP placed centromeric to DXS83 based on clone 129.

Clones 12, 52, 143 and 203 are all negative for DXS178 but score positively for DXS87, indicating that DXS87 is distal to DXS178. Likewise, clones 12 and 143 indicate that DXS87 is proximal to PLP and DXS83. The exact position of DXS87 in relation to the other probes is not clear from this data, however. Initial analysis of clone 52, which is negative for DXS178 and DXS94 but positive for DXS87, would indicate that DXS87 is situated between DXS17 and PLP. However, clone 52 is marked in Table 4.3 as having a weak positive signal with DXS87. This is in respect of the fact that the signal
produced with this probe was only one third to one half of the strength expected, based upon comparative signals obtained with other probes on the same filters. One possible explanation for this is that DXS87 is contained on a separate DNA fragment to the other probes in clone 52 and that a subpopulation of the cells that the DNA was prepared from had lost the second fragment which contained DXS87. This would necessitate the position of DXS87 as being either telomeric or centromeric to the other positively scoring probes in clone 52, as the chances of two overlapping fragments of Xq22 origin in the same clone would imply that the initial fusion had involved two C12D cells. The probability of this is extremely low as the efficiency of the original fusion was only 1 per $10^4$-$10^5$ cells (Benham et al., 1989). Therefore, this would position DXS87 proximal to DXS147 or distal to DXS83, the latter position of which is refuted by clones 12 and 143. In Section 4.1.2, it was suggested that DXS87 could either be proximal to DXS94 on the same Mlu I fragment, or alternatively DXS87 could be situated on a different fragment. RH mapping has not clarified this position.

4.3 Attempts to Physically Link DXS178 and PLP to DXS94

The results from the PFGE and IFGT hybrids indicate that, with the possible exception of DXS87, DXS94 is the closest marker telomeric to DXS178 of all the probes in this study. In an attempt to determine the physical distance between DXS178 and DXS94, a number of additional restriction enzymes were used which had some of the least frequent cutting frequencies (see Table 1.1). This was necessary because from the physical map which includes DXS94 in Figure 4.14, DXS94 is a minimum physical distance of approximately 2Mb from DXS178. The enzymes Sac II, Eag I, Not I and Nru I, which have low numbers of restriction sites in the genome (see Table 1.1), were therefore used for these digests. As can be seen from Figure 4.15, the two loci did not share common Eag I or Nru I fragments. DXS178 recognised an Eag I band of less than 1Mb whereas DXS94 recognised a 2.6Mb band. DXS94 recognised a large 3.1Mb
DXS94 and DXS178 are located on separate *Eag* I and *Nru* I Fragments

DNA from female leukocytes was digested with *Nru* I or *Eag* I and separated on a 1% gel using CHEF. Run conditions were 50V with a switch time of 2000 secs for 180 h. The buffer temperature was 11.5°C. Size markers (Mb) are *H. wingei* and *S. pombe* chromosomes.
Nru I band: this was known to be a partial digest because DXS94 has previously detected Nru I fragments of 510 and 350kb (data not shown). With the partial Nru I digest (Figure 4.15), DXS178 did not hybridise to a discrete fragment of DNA but instead to a smear. Unfortunately, only LM hybridisation was detected with both probes when DNA digested with Not I was resolved on gels with resolution of over 5.7Mb (data not shown). Natural partial Mlu I digests also failed to link DXS94 to DXS178. Mlu I often produces natural partial digests because a high proportion of potential Mlu I restriction sites occur in bulk genomic DNA instead of CpG islands and therefore a large proportion of each restriction site in a genomic DNA preparation is likely to be methylated (see Table 1.1). If the two Mlu I fragments recognised by DXS178 and DXS94 were adjacent, it may have been possible to detect a partial digestion product of slightly under 3Mb. Instead, DXS178 gave a partial product in the range of 1.5Mb (see Figure 4.9). Sac II was not useful in linking these two probes either, as DXS178 hybridised to a small 150kb band (see Figure 5.1).

In Section 4.2, radiation hybrid mapping placed the PLP-DXS83 linkage group distal to the DXS94 group, with PLP most centromeric. PLP was therefore hybridised to the filter on which DXS94 recognised large fragments with Eag I and Nru I (Figure 4.15). No similar bands were detected with PLP, which recognised an Eag I fragment of 1.1Mb and a 1.7Mb Nru I fragment.

4.4 Examination of the Telomeric Deletion Boundary of MBU by PFGE

The closest flanking marker centromeric to XLA from linkage studies is DXS3 (see Section 1.4.1.1). It has already been described in Section 1.4.1.3 that the TCD patient MBU has a deletion extending telomeric to this probe, and therefore that this deletion boundary forms the centromeric limit for the XLA locus.

In Section 3.1, it was demonstrated that the MBU deletion did not extend distally to DXS178. In an attempt to determine how far away the MBU deletion breakpoint was from DXS178, DNA from MBU was mapped using PFGE.
DXS178 was hybridised to filters containing DNA from MBU and 2X leukocytes which had been digested with the restriction enzymes Eag I or BssH II. If the MBU deletion boundary was between DXS178 and the enzyme sites, then an altered band size would be expected in the MBU tracks compared to normal controls,

\[ \text{Control DNA } E \quad \boxed{} \quad E \]
\[ \text{MBU DNA } E \quad \boxed{} \quad E \]

where \( \boxed{} \) represents DXS178, E represents enzyme restriction sites and \( == = \) represents the DNA from the far side of the deletion in MBU. The results can be seen in Figure 4.16 and in Table 4.4.

The MBU cell line showed altered band sizes with both BssH II and Eag I. The BssH II band migrated only slightly further in MBU than in the control track and this by itself could be explained simply as a DNA loading effect. The Eag I band sizes were markedly different, however, indicating that Eag I was digesting the DNA at a new restriction site in MBU.

### 4.5 Screening XLA Patients for Large Deletions close to DXS178

Whilst the majority of genetic disorders arise as a result of point mutations, a small proportion arise as the result of chromosome deletions ranging in size from a few bp to thousands of kb of DNA. The identification of a deletion in an XLA patient would be extremely useful as it would form a clear focus for cloning attempts either by the PERT technique or by chromosome walking with cosmids or YACS, depending on the size of the deletion.

It is extremely difficult to make EBV-transformed B-cell lines from peripheral blood samples from XLA patients because of the almost complete absence of a circulating B cell population. However, it is possible to produce large numbers of T cell blasts using the mitogen PHA in conjunction with the cytokine IL2. DNA plugs were made from several unrelated individuals from XLA pedigrees using T-cells produced in this manner by M. de Weers.
DNA from female leukocytes (F) and MBU was digested with *Bss* II and *Eag* I and separated on a 1% gel using CHEF. Run conditions were 120V with a switch time of 120 secs for 36 hours. The buffer temperature was 15°C. Size markers (kb) are YP148 chromosomes.
<table>
<thead>
<tr>
<th>Restriction enzyme</th>
<th>Band size (kb)</th>
<th>MBU</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eag</em> I</td>
<td>330</td>
<td>530</td>
<td></td>
</tr>
<tr>
<td><em>BssH</em> II</td>
<td>460</td>
<td>490</td>
<td></td>
</tr>
<tr>
<td><em>Sac</em> II</td>
<td>150</td>
<td>150</td>
<td></td>
</tr>
<tr>
<td><em>Sac</em> II/ <em>BssH</em> II</td>
<td>80</td>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>

The bands obtained with *Eag* I and *BssH* II are shown in Figure 4.16.
linked nature of the disease has already been established for all the pedigrees (Mensink et al., 1986a). Patient 66-8 322 is a female whose carrier status has been determined by X chromosome inactivation studies (M de Weers, personal communication). All of the other individuals are affected males.

We used PFGE to examine these DNA samples for large deletions. The marker used was DXS178, because it is the closest linked probe to XLA. The enzyme used to cleave the DNA was BssH II. This was chosen because it produced one of the largest DNA fragments detected by DXS178 ie. approx 450kb. In order to control for DNA loading effects, the filters were also hybridised with DXS94 which gives a band in a similar size range. The autoradiographs for DXS178 and DXS94 are shown in Figure 4.17. It can be seen that DXS178 recognised a band in all patients of approximately 450kb. There is track to track variation in the distance migrated by this band, however. For example, in the DXS178 autoradiograph, the DNA from patients B-2 25 and F RRA have migrated further than the DNA from H236 and M-3 76. This could be explained either by DNA loading effects or by a small deletion within the 450kb BssH II fragment recognised by DXS178. The marker DXS94 was hybridised to the same filter and a band of approximately 380kb was seen in all tracks. Again, the distance that the band migrated was found to vary slightly from track to track. Comparison of the two autoradiographs revealed that in five out of six cases, the difference in migration between the two bands was consistent for each track. The internal control for this gel was the correct-sized band in patient 66-8 322, which must have represented her normal X-chromosome. The sensitivity of this technique was estimated to be sufficient to detect deletions which were 20kb or larger in size, as this would have resulted in a 2-3 mm difference in the migration of the two bands. This indicated that in none of these families was the XLA disorder caused by a deletion greater than 20kb within a 450kb region surrounding DXS178. With patient 66-8 322 there was a small chance that a BssH II site itself had been deleted and that the aberrant chromosome was represented by a large fragment that had not migrated
DNA from PHA-stimulated T cell blasts was digested with Bss H II and separated using CHEF on a 1% gel. The switch time was 70 secs and the gel was run at 170 V for 40 h with a buffer temperature of 14°C. Size markers (kb) are YP148 chromosomes.
4.6 Discussion

**PFGE Mapping of Xq22**

Mapping probes in Xq22 by PFGE and by the use of radiation hybrids has produced valuable information regarding the organisation of this part of the X chromosome, which previously has not been mapped in great detail.

Firstly, PFGE has made it possible to physically link several probes in Xq22 and determine their relative positions. In all cases of physical linkage except one ie. that of DXS54 to DXS83 and DXS24, physical linkage has been demonstrated on at least two restriction fragments.

**Comparison of Physical Distance and Recombination Frequency in Xq22**

PFGE has made it possible to directly compare genetic recombination frequency and physical distance for a pair of markers in Xq22. As discussed previously, the convention of 1Mb = 1cM is only approximate and severalfold variations in this relationship can be found throughout the genome. In Figure 4.14, a physical map of part of Xq22 which incorporates DXS94 and DXS17 is shown. From this map, DXS94 and DXS17 are a minimum of 0.6Mb and a maximum of 0.9Mb apart. Two-point linkage data for this pair of probes is available from two studies. Arveiler *et al.* (1987) found a maximum lod score of 13.03 at Θ=0.02, with 95% confidence intervals of 0.0001 < Θ < 0.09. More recently, Kwan *et al.* (1990) obtained a maximum lod score of Z=3.42 at
\( \Theta = 0 \), with 95% confidence intervals of \( 0 < \Theta < 0.17 \). Using the Haldane mapping function, the confidence intervals for the larger study convert into genetic distances of 0.0001 and 9.9cM. The physical distance obtained in this thesis are compatible with these limits.

In addition, the map in Figure 4.14 indicates that DXS178 cannot be less than 2Mb from DXS94 as these two markers would be located on the same \textit{Mlu I} fragment if this was so. In Arveiler \textit{et al.} (1987) the two-point linkage data for this pair of loci gives a maximum lod score of \( Z = 10.31 \) at \( \Theta = 0 \), with 95% confidence limits of \( 0 < \Theta < 0.06 \). Kwan \textit{et al.} (1990) obtained \( Z_{\text{max}} \) at \( \Theta = 0.15 \), with 95% confidence limits of \( 0.03 < \Theta < 0.36 \). A minimum physical distance of 2Mb is therefore compatible with the genetic data.

\textbf{IFGT Hybrid Mapping}

The classic method of interpreting radiation hybrid data is by statistical analysis. This is based on the observation that the closer two DNA markers are to each other, the higher is the probability that they will cosegregate after random radiation-induced chromosome breakage.

The number of hybrids analysed in Section 4.3 was too small for statistical analysis and only included positively selected hybrids. However, previous analysis of seven clones from the same series of IFGT hybrids with sets of closely linked probes showed that all seven hybrids maintained the linear integrity of their inserted DNA (Benham \textit{et al.}, 1989). Benham \textit{et al.} (1989) also demonstrated that in the hybrids used in this study, the size of each human fragment, based on the number of adjacent probes found in the clones, was estimated to range from \( 10^6 \) bp to \( 10^7 \) bp. With the exception of alphoid
centromeric sequences, there did not appear to be selection for specific DNA fragments; however several hybrids had more than one insertion.

One theoretical problem that could have arisen in the present study, therefore, was that a single hybrid could contain one human fragment from proximal Xq22 and a second, non-overlapping fragment from distal Xq22. In such a hybrid, a probe from central Xq22 could score negatively and be positioned wrongly in a proximal or distal flanking position. It has already been mentioned that the eight hybrids were chosen on the basis of a primary screening panel. The data from this panel was therefore used to calculate the probability of a spurious lineage result caused by the insertion of a second fragment. The scores for the hybrids on the primary screening panel are shown in Appendix A2. On average, the clones scored positive for four out of 23 probes (excluding DXS17, DXYS5 and DXS11). This indicated that the chances of any one DXS17-positive clone scoring positive for an unlinked probe on the basis that it contained a second fragment was 21%. This data was in fact skewed by one particular hybrid (No. 52) which scored positive for 11 out of 23 probes. For the remaining seven hybrids, the chances of there being a second unlinked fragment were approximately 17%, corresponding to between 1 and 2 hybrids. This is probably an overestimation of the risk of a second close-by fragment being inherited. One would have expected Benham et al. (1989) to have one or two hybrids without complete linear integrity over short distances in their sample analysis, based upon these calculations. However, in the case where the radiation hybrids have enabled the orientation of the two large clusters of probes and also the orientation of the PLP-DXS83 group, only one or two hybrids have been informative in each case. This indicates that the RH data should be used as a guide to indicate which probes it would be provident to attempt to link together next by PFGE in order to confirm the probe orders that are implicated form the RH data.

The approximate size of the human Xq22 fragment in each hybrid was
estimated by adding together the size of the Mlu I fragments corresponding to the probes that scored positive in each hybrid. The average score was was 4-5Mb, which is in agreement with earlier estimates (Benham et al., 1989).

The use of PFGE and IFGT hybrids as complimentary mapping techniques has made it possible to order nine probes in Xq22. This order is summarised in Figure 4.18.

**Attempts to link DXS178 to DXS94**

One of the original aims of this study was to produce a continuous physical map of the XLA locus which incorporated the locus boundaries. It is likely that DXS178 has been physically linked to one boundary i.e. the MBU deletion breakpoint. The use of a large range of infrequently cutting restriction enzymes and partial enzyme digests with PFGE conditions designed to separate large DNA fragments (>3Mb) has not achieved the physical linkage of DXS178 and DXS94, however. Mapping data from the PFGE and from IFGT hybrid analysis indicates that of all the probes used in this study, none with the possible exception of DXS87 would be useful in bridging the distance between DXS178 and DXS94. The ascertainment of a precise physical value for this distance is important as it will influence the type of approach taken to clone the XLA locus. On the basis of the data obtained during this study, and upon linkage data, this distance may be as small as 2-5Mb. If, after further investigations, the locus size is at the lower end of this estimation, then it will be feasible to obtain YAC clones to cover this distance and perhaps to use them to screen B cell cDNA libraries and isolate cDNAs which would be candidates for the XLA gene.

**Estimation of the Physical Size of Xq22**

The minimum distance mapped in Xq22 can be calculated by adding the sizes of non-overlapping Mlu I fragments detected in 2X DNA together. Adding together 1.5Mb (partial) for DXS178, 2.7Mb for the DXS94-DXS17 region,
Figure 4.18
A Map of Xq22: a Summary of PFGE and Hybrid Mapping Data

Key: E=Eag I, M=Mlu I, S=Sfi I
590kb for PLP and 1060kb for DXS83-DXS54 gives a total of 6Mb. This rises to 8.5Mb if DXS87 is on a separate Mlu I fragment. A very approximate estimate of Xq22, made by measuring the Xq22 locus on a high resolution chromosome map (Ferguson-Smith, 1974) gives a locus size of 10-12Mb. This suggests that at least 50-60% of the Xq22 locus has been covered by pulsed field mapping.

By examining the G banding pattern and comparing this to the summary map in Figure 4.18, it could be postulated that DXS178 is situated in the reverse band Xq22.1; DXS17 and linked probes are in the G dark band Xq22.2 and PLP is in the reverse band Xq22.3. Gardiner et al. (1990) found a good correlation between G-bands and the size of rare-cutting enzyme fragments found: light bands were associated with smaller fragments and dark with larger fragments. This is substantiated by looking at the sizes of the Mlu I and BssH II bands obtained by the different probes in Xq22.

Comparison of Murine and Human Chromosomes

In Section 1.5, the observation was discussed that the X chromosomes of all mammals appear to be syntenic, and that this phenomenon extends to homologous subchromosomal regions in the mouse and the human which frequently contain conserved gene order in addition (see Figure 1.4). If conserved linkages can be shown within these homologous regions then it aids the prediction of the position of a particular gene in one species if the position of that gene has already been determined in the second species. One of the syntenic regions described in Section 1.5 contains the loci EDA (Ta), PGK (PgkI), PLP (jp) and GLA (Ags). Also contained in this group is xid in the mouse and XLA in the human (Figure 1.4). Previously, it has been unclear whether or not this syntenic group also showed conserved linkage. The use of interspecies crosses to generate highly informative meioses with polymorphic probes has enabled the high resolution ordering of probes from this region in the mouse (Amar et al., 1988). This information, combined with the positioning
of xid (Berning et al., 1980) has made it possible to determine the order of murine loci in this region as shown in Figure 1.4., i.e. centromere-Ta-Pgkl-xid-Ags-Plp.

Information regarding the ordering of the equivalent loci on the human X chromosome has been less detailed, however. The human gene order EDA-PGK1-(GLA, PLP) can be extrapolated from the by mapping data of Cremers et al. (1988) and appears to be equivalent to the order of murine loci. From the mapping data obtained in this study, PLP is distal to DXS94 and therefore distal to XLA too. Astrin et al. (1989) have reported that linkage studies place GLA distal to DXS17. As DXS17 flanks XLA, GLA can also be positioned distal to XLA. This means that the human gene order is EDA-PGK1-XLA-(GLA, PLP), which strongly indicates a conserved linkage group between the two species and therefore an increased likelihood that xid is the murine homologue for XLA on positional grounds. A more detailed look at this region indicates that there are discrepancies in the distal halves of these regions, however. The mapping data in Section 4 of this study clearly indicates that in the human, the marker p212 (DXS178) is proximal to PLP. This gives a locus order in the human as PGK1-(212, XLA)-(PLP, XGAL) whereas the murine order is Pgkl-xid-(Ags-cX52.5)-Plp-p212. This discrepancy in the order of p212 indicates that additional chromosome rearrangement events may have occurred in the distal section of these equivalent regions. Unfortunately, the comparative order of the loci PLP, GLA and cx52.5 in the human have not yet been elucidated. One single rearrangement event that could account for the differential gene orders in the murine and human regions would be the inversion of the distal end of the region (see Figure 4.19). If, in this speculative scheme, XLA and xid are homologous, then the XLA gene would be proximal to p212. If it were distal, then the position of xid in the mouse would be more distal following the divergent evolution of the two chromosomes. At present this scheme is only speculative. However, the PFGE maps of Xq22 and the radiation hybrid panel that have been
Figure 4.19
Comparison of Homologous Regions on the Human and Murine Chromosomes Containing the XLA/xid Genes

See text for explanation of probe order. Brackets ( ) indicate that relative order is not known. The arrows indicate the site of a possible inversion event that may have occurred within this subchromosomal region during the divergent evolution of the human or murine chromosomes.
established during this study will facilitate the determination of the relative orders of GLA, PLP and cX52.5. If the order of the human markers is found to be the exact inverse of the murine homologues, then the possibility of a single additional inversion event would have more credence. If this proves to be the case, it may be prudent to closely search the area proximal to DXS178 for the XLA gene. This area could be as small as 330kb if the MBU deletion data is substantiated.

**Mapping the MBU deletion breakpoint**

The presence of a new *Eag* I restriction site in MBU can be explained in three ways, which are:

1) The deletion breakpoint in MBU is situated between DXS178 and the nearest proximal restriction site which can be digested by *Eag* I. The smaller band in MBU is due to the existence of an *Eag* I site in the DNA on the proximal side of the MBU deletion and this new *Eag* I fragment spans the deletion.

2) The MBU lymphoblastoid cell line has become demethylated, resulting in the availability of a previously methylated *Eag* I site for digestion.

3) MBU has an *Eag* I polymorphism.

Option 2 relies on the assumption that the *Eag* I site in the MBU cell line is present but methylated in the control DNA. However, statistically, 75% of all *Eag* I sites are located in CpG islands and therefore are likely to be unmethylated on the active X chromosome of the 2X control DNA. Additional work with DXS178 and the MBU EBV line showed that this marker recognised similar sized bands to 2X control DNA with either *Sac* II digests (140kb) or with *Sac* II/BssH II double digests (75kb), indicating firstly that the deletion did not extend to within the *Sac* II sites surrounding DXS178 and also that MBU generally showed similar methylation patterns to non-transformed leukocytes. If option 1 is correct and the band shift in MBU is due to the loss of the original *Eag* I site and replacement with a new site from the centromeric flank of his deletion, then the deletion boundary, which defines the centromeric limit for the
XLA locus, must commence within 330kb of DXS178.

In order to confirm this, a number of additional EBV-transformed lines should be digested with *Eag* I and run in parallel with MBU, which should indicate whether or not such cell lines tend to gain or lose methylation upon culture. In Section 4 of this study, the cell lines LCL-127 and AnLy both showed differences in methylation of *Mlu* I sites compared to 2X peripheral leukocytes; this was detected by the markers PLP, DXS24 and DXS54 (Figure 4.3). The majority of *Mlu* I sites are situated outside CpG islands, whereas the opposite is true for *Eag* I (see Table 1.1). It is not known whether or not the occurrence of restriction sites in particular CpG islands compared to those in genomic DNA has any effect on the maintenance of methylation status in culture. The enzyme *BssH* II, whose site distribution is similar to that of *Eag* I, was used in studies with PLP, DXS54 and DXS24 (Figure 4.5). The band sizes obtained for all three markers were identical to those obtained with the 2X leukocytes.

Option 3 is extremely difficult to test, although partial *Eag* I digests with MBU should reveal the larger 530kb band in addition to the smaller 330kb band. In addition, examination of at least 10 unrelated individuals with DXS178 and *Eag* I has not revealed any frequent polymorphism of this site (M. de Weers, personal communication).

Option 1 therefore seems the best explanation for the altered band in MBU DNA. The only way to determine this with certainty is to examine MBU DNA with a marker such as DXS121 (p784) which has been shown to detect a 430kb *Sal* I fragment and a 400kb *Sfi* I fragment which both span the centromeric breakpoint of the MBU deletion and which, by double digest studies has been shown to lie within 200kb of the breakpoint (Cremers *et al.*, 1989; Merry *et al.*, 1989). By choosing an appropriate restriction enzyme, both DXS178 and DXS121 should be able to detect the same DNA fragment in MBU DNA. *Eag* I or *BssH* II may provide an appropriate digest for this, although the small size
(330kb) of the MBU Eag I band indicates that the new Eag I restriction site is close to the deletion boundary. As will be discussed in Section 5, many of the rare-cutting enzymes produce small digest products with DXS178, possibly because of the presence of CpG islands, therefore it may well be necessary to try some different enzymes, such as Pac I (New England Biolabs) which is an eight-base cutter but whose sites are not associated with CpG islands. Mlu I may also be useful, as DXS178 gives a large partial digestion product with this enzyme.

Assuming that Option 1 is correct, the size of the XLA locus has been reduced significantly. Instead of being 10cM (the genetic distance between the flanking markers DXS3 and DXS94, it is now 5cM (the genetic distance between DXS178 and DXS94) plus the maximum 330kb between DXS178 and the MBU deletion boundary.

**Examination of XLA Patients' DNA using PFGE**

The success of positional cloning approaches rely heavily on reducing the potential size of a gene locus to a size small enough to handle as cloned segments. This can occasionally be accomplished when extremely good linkage data and physical mapping data is available, as was the case with the CF gene locus (Rommens et al., 1989b). Linkage data as informative as this is extremely difficult to obtain in disorders as rare as XLA, however, as it is not possible to study a sufficiently large number of meioses. The majority of successful positional cloning projects have managed to define the disease locus by the detection of rare patients with large deletions. The proportion of genetic disorders which are caused by gross chromosomal deletions, as opposed to point mutations, varies widely between genetic disorders. For example, in DMD, approximately 70% of new mutations appear to involve large deletions (den Dunnen et al., 1987; Forrest et al., 1988). One might expect a high deletion rate for DMD because of the extremely large size of the gene (2000Mb). The incidence of deletions in other diseases indicates that factors other than gene size influence the deletion rate, however. Presumably, the positioning any genes
whose dysfunction is lethal in utero will affect the position of male viable deletions on the X chromosome. The factors which determine that the majority of X chromosome deletions are restricted to only two G- dark bands (Section 1.4.4) seem more complicated than size and position, however. For example, the CF gene on chromosome 7 spans approximately 250kb of genomic DNA, yet no gross deletions have been recorded to date. In contrast, the HPRT gene is only 1/6th of the size of CF gene yet is deleted in up to 10% of Lesch-Nyhan patients (Yang et al., 1984). The STS gene in Xp21.3 spans 150kb of genomic DNA, yet up to 90% of patients have large deletions averaging approximately 2 Mb. A model for this has recently been proposed where examination of the junction fragments of such deletions indicates that both breakpoints are situated within a family of homologous fragments (Yen et al., 1990). Research into the X-linked immunodeficiency XLP has recently resulted in the detection of a cytogenetically detectable deletion in one out of 59 families (Purtilo and Grierson, 1991). Therefore it seems difficult to predict whether an XLA patient with a large deletion will be found or not. In order to increase the chances of finding a deletion in an XLA patient, the strategies must be firstly to maximise the number of unrelated XLA patients that are examined by PFGE. Secondly, the range of enzyme/DNA marker combinations used to examine each DNA sample should optimise coverage of the potential XLA locus. One such combination which will cover the most distal 2Mb of the locus is Mlu I and DXS94 (see Figure 4.13). The PFGE run conditions used to separate such large fragments tend to give a final resolution of 1-2mm of gel per 100kb DNA, so small changes in fragment size would not be visible, however any large deletion would be detected. Mlu I could also be used at low concentrations to produce reliable partial digest fragments (approx 1.5Mb) in combination with DXS178 (See Figure 4.9). These digests could be resolved more clearly by using PFGE with different running conditions.
5.1 Detailed Physical Mapping around DXS178 using PFGE

Several of the infrequently cutting restriction enzymes used in this study gave quite small DNA fragments (less than 150kb) upon hybridisation with DXS178. This had two major implications. The first was that it made it extremely difficult to physically link DXS178 to the nearest markers (ie. DXS94 and the MBU deletion boundary). The second was that the presence of several restriction sites for infrequently cutting restriction enzymes could indicate that there were CpG islands close to DXS178. In order to determine whether or not these sites were clustered, a series of single and double digests were carried out with these enzymes.

Female leukocytes all have one activated and one inactivated X chromosome. CpG islands are frequently methylated on inactivated X chromosomes. If the enzyme sites under investigation were located in CpG islands then there was a possibility that female leukocyte DNA (Figure 5.1a and 5.1b) would produce a complex series of results. For this reason, several of the digests were repeated using DNA from male leukocytes (Figure 5.1c) with the rationale that all X chromosomes should be activated and therefore in the same state of methylation. The filters were all hybridised with DXS178. The pattern of bands was found to be similar with both male and female leukocyte DNA. The band sizes obtained with any one restriction enzyme varied from filter to filter by approximately 10kb (See Figure 5.1) so similar enzyme digests were used as internal controls when bands from separate filters were compared. The fragments sizes obtained are listed in Table 5.1.

From this PFGE data, a long-range physical map of the region surrounding DXS178 was constructed (see Figure 5.2). It can be seen that the enzyme sites are not evenly dispersed, but fall mainly into three or four clusters which have been numbered 1 to 4 in Figure 5.2. The coincidence of several rare-cutting
Figure 5.1
PFGE Mapping around the DXS178 Locus

Key:
M = Mlu I
B = Bss H II
S = Sac II
F = Sfi I
N = Nar I
A = Nae I
L = Sal I
E = Eag I

See Figure 5.1 c for legend.
DNA from female (a and b) and male (c) leukocytes was separated on a 1% agarose gel using CHEF. 5.1. (a) was run for 20 h using a switch time of 5 secs, a voltage of 190V and a buffer temp of 11 °C; 5.1 (b) was run for 15 h with a switch time of 7 secs. 5.1 (c) was run for 20 h with a switch time of 7 secs, at 190V and 13 °C. Size markers (kb) are bacteriophage λ concatamers.
Table 5.1

Restriction Fragments Obtained from Genomic DNA with DXS178

<table>
<thead>
<tr>
<th>Restriction Enzyme(s)</th>
<th>$Bss$HII</th>
<th>Sfi I</th>
<th>$Sfi$ I</th>
<th>Sac II</th>
<th>Sac II</th>
<th>Sac II</th>
<th>Mlu I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment Size (kb)</td>
<td>470</td>
<td>135</td>
<td>125</td>
<td>150</td>
<td>80</td>
<td>90</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>90(weak)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Restriction Enzyme(s)</th>
<th>$Mlu$ I</th>
<th>$Sac$ II</th>
<th>Nae I</th>
<th>Nae I</th>
<th>Nar I</th>
<th>Nar I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment Size (kb)</td>
<td>80</td>
<td>150</td>
<td>155</td>
<td>155</td>
<td>&gt;200</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>80</td>
<td></td>
<td>78</td>
<td></td>
</tr>
</tbody>
</table>

The band sizes contained within this table were taken from the three blots in Figure 5.1. The maximum variation in band size over all three gels was 15 kb. The Sac II digestes were used to standardise sizes between the different blots.
Figure 5.2
A Physical Map Showing Unmethylated Enzyme Sites around the DXS178 Locus

Key: M = Mlu I, B = BssH II, S = Sac II, F = Sfi I, A = Nae I, N = Nar I.
1, 2, 3 and 4 indicate clusters of restriction sites.
( ) = only partial digestion seen
restriction sites such as the ones in Figure 5.2 is strongly indicative of the presence of unmethylated CpG islands. It has already been illustrated in Table 1.1 that the distribution of restriction sites for the enzymes used in Figure 5.2 varies within and outside CpG islands. Indeed, the number and range of sites that have been mapped around DXS178 varies considerably between the CpG-rich regions. However, from the PFGE data alone it is not possible to tell how many times each enzyme cuts at one particular CpG-rich region. In addition, it is not always possible to determine whether enzymes that have sites in regions 2 and 3 also have sites in regions 1 and 4. The proportion of DNA that has been digested at particular sites varies too. For example, the $Mlu$ I and $SfI$ I sites at cluster 2 often do not digest completely. This could reflect the presence of a gene which is being transcribed in a proportion of cells from which the DNA was extracted (for example, lymphocytes).

5.2 Isolation and Mapping of DXS178 Positive Cosmids

It was desirable to screen a cosmid library with DXS178 for two reasons. Firstly, the CpG islands identified in the region surrounding DXS178 are possible indicators of transcribed sequences. The close linkage between XLA and DXS178 means that such sequences should be considered to be possible candidates for the XLA gene. Secondly, a chromosome walk past such islands would allow the isolation of new DNA probes which could be used on PF gels to determine the distance from DXS178 to DXS94. Information gained in this manner may affect the approach that should be taken next towards cloning the XLA gene as previously discussed.

5.2.1 Isolation of a Single-Copy Fragment of DXS178

The DNA marker 212/9 which detects the locus DXS178 is a mixture of single-copy and repetitive sequences and requires preannealing before hybridisation to Southern blots. In order to screen a cosmid library for DXS178
positive clones, it was necessary to isolate a fragment from p212/9 that was free of repetitive DNA. This was done by hybridising $^{32}$P-labelled sonicated placental DNA to a filter containing different enzyme digests of the 212/9 plasmid and determining which fragments of 212 hybridised poorly with the probe, which effectively contained repetitive DNA. This experiment indicated that the bulk of the repetitive sequences were located in a 2.3kb EcoR I/Taq I fragment whereas a 7kb EcoR I/Taq I fragment was probably single-copy as only a very faint hybridisation signal was obtained (data not shown). The 7kb fragment has been designated 212A and the smaller fragment 212B in this study.

5.2.2 Isolation of DXS178 Positive Cosmids with 212A

The screening of the cosmid library ICRFc100 with 212A has been described in the Materials and Methods section. Seventeen "positive" clones were originally picked from the library. This was a larger number of positives than expected as the library contained a 4-5 fold representation of the X chromosome (a single X chromosome is 1.5x10^8 bp and the library contains 16,000 colonies with an average insert size of 40kb). Therefore only 4-5 positive clones were expected. The intensity of the signal varied between colonies, falling into three main categories (very strong, medium and weak). This raised the possibility that one or more additional DNA sequences might exist in the genome with quite high homology to 212A. Ten cosmid clones remained after the elimination of duplicated and negative clones (see Materials and Methods) and large scale DNA preparations were made of these. The DNA was digested with Hind III or EcoR I, electrophoresed and blotted. The blots were hybridised with probes 212A, 212B and Lorist X. 212A hybridised very strongly to 4 of the clones. In 3 of these, ie. ICRFc100D1020 (D1020), ICRFc100E03167 (E03167) and ICRFc100D0590 (D0590) the positive signal in the EcoR I digests came from a band of approximately 10kb, equivalent in size to the genomic EcoR I fragment from which 212A was isolated. In a fourth clone, ICRFc100C0897 (C0897),
212A hybridised to a smaller band of approximately 7kb in size. Probe 212B detected identical bands in these four cosmids which strongly indicated that they contained genomic DNA complimentary to the original 212/9 probe. The remaining six clones showed weak hybridisation to 212A and none to 212B. The signal obtained from these clones with 212A (as estimated by eye) ranged from approximately one tenth to one quarter of the intensity produced with the other 4 clones. This suggested the possibility that the six weakly hybridising clones either contained a very small proportion of the 212A sequence and therefore overlapped the 212A+B+ clones, or contained sequences other than 212A but with enough homology to cross-hybridise with 212A. The fragment sizes obtained with the six clones appeared to be completely dissimilar to those obtained with the four strongly hybridising clones, giving no early indication that they overlapped. Therefore the mapping studies of these cosmids commenced with the clones that were known to contain the complete 212A and 212B sequences.

The sizes of the genomic inserts were estimated using PFGE. The clones were digested with BssH II, Sac II or Cla I and separated using PFGE. Clones D1020, E03167, D0590 and C0897 were linearised with Sac II and Cla I which was compatible with a single site within the vector (See Appendix A3) and no sites within the insert. The sizes of the linearised cosmids were approximately 46kb, 46kb, 48kb and 16kb respectively. All four cosmids contained a BssH II site in their genomic inserts in addition to the vector site.

5.2.3 Restriction Enzyme Mapping of Cosmids ICRFc100D1020 and ICRFc100E03167

In order to determine whether either of the CpG rich regions closest to DXS178 had been cloned in these cosmids, restriction maps of clones D1020 and E03167 were constructed. Preliminary studies showed that both these cosmids were cleaved into two fragments after digestion with Kpn I. This enzyme was therefore used to digest the cosmid DNAs in combination with a large number of
infrequently cutting restriction enzymes. The digests were electrophoresed and blotted and the filters were hybridised with the probes 212A and Lorist X. Examples of these digests are shown in Figures 5.3 and 5.4. A diagrammatic representation of Lorist 4 showing restriction sites relevant to this study can be found in Appendix A3. Several of the restriction enzymes had sites within the vector that cut obliquely in relation to the _BamH_I cloning site. This was advantageous in mapping the cosmids because the probe Lorist X frequently hybridised to two bands at different intensities on a single track and allowed the orientation of the two bands. Tables 5.2 and 5.3 show the results obtained for D1020 and E03167 respectively. From these results, restriction maps of D1020 and E03167 were constructed (Figures 5.5 and 5.6).

No genomic sites were found for _Sal_I, _Mlu_I, _Sfi_I, _Not_I, or _Nru_I in either D1020 or E03167 other than in the predicted sites within the vector. Bands of less than 3kb were not visible on these digests. Additional small _Kpn_I fragments of approximately 1.7kb and 1.5kb were found on separate _Kpn_I digests for D1020. In E03167, the small _Kpn_I fragments were 1.8kb, 1.7kb and 1.5kb. The 1.7kb and 1.5kb fragments were purified and hybridised to cosmid _BssH_II/_Kpn_I digests to determine their orientation. The 1.8kb band was not present in D1020 and therefore had to be in the position shown in Figure 5.6.

Clones E03167 and D1020 are almost identical in size. Addition of the _Nae_I fragments gives a total size of approximately 46kb for both D1020 and E03167 which agree well with the PF estimations. Both inserts are in the same orientation. D1020 contains approximately 1kb in addition at the 5' end of the insert and E03167 contains an additional 1.2kb at the 3' end of the insert.

5.3 Comparison of Genomic and Cosmid Maps around DXS178

In an attempt to orientate the cosmids D1020 (Figure 5.5) and E03167 (Figure 5.6) with respect to the genomic map around the DXS178 locus (Figure
Figure 5.3
Digestion of ICRFc100D1020 with Infrequently Cutting Restriction Enzymes

DNA from a CsCl preparation of D1020 was digested with a variety of infrequently cutting restriction enzymes, either singly or in combination with KpnI (K). The digests are shown in a). Digests were Southern blotted and hybridised with b) 212A and c) Lorist X. Size markers are λ/Hind III and λ/Sal I.
DNA from a CsCl preparation of E03167 was digested with a variety of infrequently cutting restriction enzymes, either singly or in combination with Kpn I (K). The digests are shown in a). Digests were Southern blotted and hybridised with b) 212A and c) Lorist X. Size markers are λ/Hind III and λ/Sal I.
Table 5.2
Restriction Mapping and Hybridisation Data for ICRFc100D1020

<table>
<thead>
<tr>
<th>Restriction Enzymes</th>
<th>Fragment size (kb)</th>
<th>Hybridisation Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kpn I</td>
<td>&gt;28</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>▼▼</td>
</tr>
<tr>
<td>BssH II</td>
<td>28</td>
<td>*▼</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>?▼</td>
</tr>
<tr>
<td>BssH II / Kpn I</td>
<td>20</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.2</td>
<td>▼</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>▼▼</td>
</tr>
<tr>
<td>Nar I</td>
<td>&gt;28</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>13-14</td>
<td>▼▼</td>
</tr>
<tr>
<td>Nar I / Kpn I</td>
<td>&gt;28</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>7.8</td>
<td>▼</td>
</tr>
<tr>
<td></td>
<td>4.6</td>
<td>▼▼</td>
</tr>
<tr>
<td>Nae I</td>
<td>19</td>
<td>▼</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>▼▼</td>
</tr>
<tr>
<td>Nae I / Kpn I</td>
<td>&gt;28(p)</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>19(p)</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.0</td>
<td>▼</td>
</tr>
<tr>
<td></td>
<td>5.3</td>
<td>▼▼▼ not resolved</td>
</tr>
<tr>
<td></td>
<td>?5 (predicted)</td>
<td></td>
</tr>
</tbody>
</table>

Key:
▼ = hybridises with lorist X
* = hybridises with 212A
p = partial digest
Table 5.3
Restriction Mapping and Hybridisation Data for ICRFc100EO3167

<table>
<thead>
<tr>
<th>Restriction Enzymes</th>
<th>Fragment size (kb)</th>
<th>Hybridisation Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kpn I</td>
<td>&gt;28</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>▼▼</td>
</tr>
<tr>
<td>BssH II</td>
<td>28</td>
<td>*▼▼▼▼</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>▼</td>
</tr>
<tr>
<td>BssH II / Kpn I</td>
<td>20</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>11.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.8</td>
<td>▼▼ [not]</td>
</tr>
<tr>
<td></td>
<td>5.6</td>
<td>▼▼ [resolved]</td>
</tr>
<tr>
<td>Nar I</td>
<td>&gt;28</td>
<td>*▼▼</td>
</tr>
<tr>
<td></td>
<td>15-16</td>
<td></td>
</tr>
<tr>
<td>Nar I / Kpn I</td>
<td>&gt;28</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>6.8</td>
<td>▼</td>
</tr>
<tr>
<td></td>
<td>5.8</td>
<td>▼▼</td>
</tr>
<tr>
<td>Nae I</td>
<td>18</td>
<td>▼</td>
</tr>
<tr>
<td></td>
<td>14(x2)</td>
<td>*▼▼▼</td>
</tr>
<tr>
<td>Nae I / Kpn I</td>
<td>19(p)</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.6</td>
<td>▼▼</td>
</tr>
<tr>
<td></td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>?5(predicted)</td>
<td></td>
</tr>
</tbody>
</table>

Key:
▼ = hybridises with lorist X
* = hybridises with 212A
p = partial digest
5.2), the positions of the infrequently cutting restriction enzymes within the maps were compared. Two similarities were apparent. The first was the coincident occurrence of the BssH II and Nae I sites in both the cosmid maps and in region 3 of Figure 5.2. However, in the genomic map, there was an Sfi I site within 10kb of the BssH II and Nae I sites whereas there were no Sfi I sites in either of the cosmids. D1020 contained at least 18kb of insert DNA on either side of the coincident BssH II and NaeI sites, which is more than sufficient to allow for any inaccuracies on the genomic map. This indicated, therefore, that the two BssH II sites were dissimilar. There are no additional BssH II and Nae I sites in similar proximity to DXS178 on the genomic map and the sites in the cosmids are therefore most likely to be restriction sites which are methylated in genomic DNA and therefore unavailable for digestion. The majority of cloned DNA is unmethylated by the bacterial host and therefore these sites would have been made available for digestion after propagation in bacteria. A second explanation for the absence of the Sfi I site in the cosmid DNA was that it had become methylated by the dam or dcm methylase enzymes of the ED8767 host bacteria. It was not possible to predict whether the Sfi I recognition site could be methylated because the sequence is CCCCNNNNNNGGGG and the internal nucleotides may have formed a recognition site for either of these methylases. In order to test this theory, the Kpn I restriction sites either side of 212A were utilised (see Figures 5.5 and 5.6). These produced a 212A positive band in the cosmids of about 32kb, which was reduced to 20kb when the DNA was additionally digested with BssH II (Tables 5.2 and 5.3). The same effect would be expected in genomic DNA if the two BssH II sites matched. Therefore, genomic DNA was digested either singly with Kpn I or doubly with Kpn I and BssH II, separated on PFGE and hybridised with DXS178 (Figure 5.7). DXS178 recognised a 34kb band in the Kpn I track, which was very similar in size to the 32kb Kpn I fragment detected by DXS178 in the cosmids. Additional digestion with BssH II did not reduce the size of this fragment in the genomic digests, however, indicating that the BssH II site in the cosmid was
Size of entire cosmid (including vector) is 45 kb.

\[ \text{Bam} \text{H I cloning site} \]
Figure 5.6
Restriction Map of ICRFc100E03167

Size of entire cosmid (including vector) is 46 kb.

\[ \text{= Bam } H \text{ I cloning site} \]
DNA from female leukocytes was digested with *Kpn* I either singly or with *BssH* II and separated on a 1% gel using CHEF. The gel was run for 16 h at 190 V with a 3 sec switch time. Size markers (kb) are bacteriophage λ, both whole and digested with *Hind* III. Key: K=*Kpn* I, B=*BssH* II.
dissimilar to the site in region 3 on the genomic map. In Figure 5.7, there is a smeared signal in the \textit{Kpn I/Bss II} track at a lower size than the signal obtained by \textit{Kpn I} alone. Had both enzymes cut in the genomic DNA, however, the expected fragment size would have been 20 kb which is smaller than the smear which is seen.

A second possible similarity between the cosmid was the coincident occurrence of \textit{Nar I} and \textit{Nae I} sites 3' to 212A in the cosmid. A similar arrangement was found in region 2 of the genomic map. However there were also sites for \textit{Sac II} and \textit{Mlu I} in region 2 of the genomic map. In contrast, cosmid E03167 contained an insert which extended almost 10kb 3' to the \textit{Nae I} and \textit{Nar I} sites and was devoid of \textit{Mlu I} and \textit{Sac II} sites.

The most probable position of the cloned DNA was therefore between regions 2 and 3 on the genomic map (Figure 5.2). However it was not possible to precisely map the cosmids within this region nor to orientate the cosmid relative to these regions from the information available.

\textbf{5.4 Discussion}

\textit{Identification of CpG-rich clusters close to DXS178}

The identification of clusters of rare-cutting restriction enzymes close to DXS178 indicates that several CpG islands may be flanking this DNA marker. Enzymes such as \textit{Sac II} and \textit{BssH II} are frequently associated with CpG islands whereas sites for enzymes such as \textit{Nae I}, \textit{Nar I} and \textit{Mlu I} occur more frequently in bulk DNA (see Table 1.1). Approximately 90\% of CpG dinucleotides are methylated in bulk DNA, however, which greatly increases the possibility that these "islands" are genuine. During this study, a paucity of restriction sites for a number of rare-cutting enzymes within Xq22 was noted. For example, only one \textit{Not I} site was detected (with DXS101; data not shown).

The question remains as to whether or not the XLA gene is associated with a CpG island. Investigations of other tissue-specific X-linked genes show on the whole that they are not associated with CpG islands (see Section 1.7).
therefore suggests that the XLA gene will not be closely associated with the islands close to DXS178. It will of course be interesting to see what, if any, genes are encoded near these islands. It is likely that at least one will be present, because the majority of CpG islands are associated with transcribed elements. There are likely to be several important genes in this area because DXS178 is probably situated in the most proximal third of Xq22, which is a G-light band, a property frequently associated with transcribed sequences. The absence of viable males with deletions in this area also indicates that important genes could be present.

Isolation of DXS178 Positive Cosmids

Hopefully, the isolation of DXS178 positive cosmids will facilitate the construction of a complete map of the XLA locus, the size of which, as previously discussed, will determine the direction of future work in this area. The average size of the inserts in the ICRFc100 library appears to be 40kb. The position of cosmids D1020 and E03167 in relation to the genomic map containing DXS178 (Figure 5.2) indicates that the cosmid is flanked by genomic unmethylated restriction sites of Sac II and Mlu I (from region 2) and BssH II (from region 3) among others. From PFGE mapping (Figure 5.2), these regions are spaced between approximately 80kb apart. A bidirectional "step" from these cosmids should be sufficient to span the CpG rich regions 2 or 3, even if each step is only 20kb. This should enable the isolation of a probe which will recognise either flanking Mlu I and Sac II fragments on one side, or flanking BssH II fragments on the other. This may make it possible to confirm the relationship between DXS178 and the MBU deletion breakpoint or to link DXS178 to DXS94. The cloning of either the Mlu I site or the BssH II sites that flank DXS178 should make it possible to perform a chromosome "jump" and isolate probes corresponding to distant restriction sites. This method was used to extremely good effect in the cloning of the CF gene, for example (Rommens et
al., (1989a). The *Mlu* I and *BssH* II sites will be especially useful for this as jumping libraries made from a 4X cell line have already been constructed with these enzymes (A-M. Poutska, Heidelberg). Again, this approach will facilitate the completion of a complete physical map of the XLA locus.

The tight linkage of XLA and DXS178 suggest that it may be useful to screen these cosmids for sequences conserved between species using so called "Zoo" blots. DXS178 is already known to have a murine homologue. It is worth noting, however, that the *xid* and p212 loci are separate in the mouse.

The cosmids can also be labelled and hybridised in sections to Northern blots from B-lineage cells, in order to detect any sequences from the cosmids that are transcribed in these cell types.
6. DISCUSSION: SCOPE FOR FUTURE WORK

6.1 Summary

Previous to the submission of this thesis, genetic and deletion mapping studies had determined that the XLA locus was situated in Xq22 and was flanked centromerically by the breakpoint of a constitutive deletion in a choroideremia patient and telomerically by the polymorphic marker DXS94.

In the present study, genetic linkage analysis, PFGE, radiation hybrid panels and finally cosmid cloning were employed in an attempt to further refine the XLA locus and to increase the likelihood of this gene being cloned on the basis of its location.

In this study, XLA families were investigated with the probe DXS101 and no recombinations were found with the disease locus in ten informative meioses.

DNA from the choroideremia patients MBU, NP and DM was used to refine the localisations of several probes and of these, the three markers DXS83, DXS87 and DXS178 were found to map to the Xq22 region containing the XLA locus.

These probes and several other DNA markers, chosen on the basis of their chromosomal location (see Section 1.8.1 and Figure 1.6), were used to construct a physical map of Xq22, the region containing the XLA locus. The majority of probes were incorporated into a map which was in three, unlinked sections and a panel of radiation hybrid cell-lines was used to orientate the separate sections of the map relative to each other and to the centromere.

All of the additional probes which were incorporated in the physical map of Xq22 ie. PLP, DXS24, DXS54, DXS83, DXS147, DXS211 were found to be distal to DXS94, previously the closest polymorphic probe telomeric to the XLA locus (see Figure 6.1). Therefore DXS94 remained the closest flanking marker to the XLA locus at the end of this study. However, the positions of DXS87 and DXS101 were not established. This was because DXS87 shared only one band
Figure 6.1
Summary: A Physical and Genetic Map of Xq21.3-Xq22

Genetic Map

- 0.001<θ<0.11
- 0<θ<0.06
- 0<θ<0.03
- 0.0001<θ<0.09

 DXS3
 DXS178
 DXS17
 DXS94

 XLA

 DXS101

Physical Map

 DXS3 deletion boundary
 MBU

 DXS178

 XLA

 DXS211
 DXS147
 PLP
 DXS54
 DXS24

 S, S
 M

 0 1
 Scale (Mb)

centromere
 telomere

Key: E=Eag I, M=Mlu I, S=Sfi I. The recombination fractions on the genetic map show the 95% confidence limits which have been calculated from Arveiler et al., 1987. The confidence limits for XLA to DXS178 have been calculated from the cumulative lod score to date.
in common with Xq22 probes on PFGE and its location could not be clarified using the X chromosome radiation hybrids; also DXS101 was found to recognise a number of dispersed loci within Xq22 in preliminary PFGE studies.

With the exception of CF, all successful attempts to isolate genes using the positional cloning approach have relied upon the identification of one or more patients with translocations or gross chromosomal deletions. Previously, no XLA patients have been found to be deleted for the probes used in genetic studies, i.e. DXS3, DXS17, DXS94 or DXS178. Therefore in order to extend this work, PFGE was used to screen XLA patients using the marker most tightly linked to the XLA locus, DXS178. The enzyme BssHII was chosen for these studies because it produced one of the largest band sizes (450kb) detected with DXS178; no alterations were detected.

The physical mapping data produced enabled the relative genetic and physical distances between polymorphic markers in Xq22 to be compared and no evidence of an unusually high or low recombination rate was found. The physical data is summarised, together with current linkage data in Figure 6.1.

The probe DXS178 which is tightly linked to the XLA locus (Z = 14.48 at θ=0), generally hybridised to small DNA fragments (<200 kb) on PFGE analysis with several infrequently cutting restriction enzymes. Double enzyme digests indicated that the restriction sites were clustered, suggesting the presence of potential CpG islands close to this marker. A bidirectional chromosome walk towards these islands was initiated by the isolation and restriction mapping of DXS178-positive cosmids. The cosmids were found to be situated between two of the CpG islands close to DXS178, within 20kb of the closest of these islands.

This study has shown the efficacy of PFGE as a long-range mapping technique, but has also illustrated the reliance of this method on a sufficient number of evenly spaced DNA markers and restriction enzyme sites in order for it to be used to the best advantage. The data produced by combining the PFGE
and RH results indicates that with the exceptions of the unplaced DXS87 and DXS101, the additional probes used in this study which map to Xq22 will not be useful in determining the physical distance between the markers which flank the XLA locus.

Although this study has not resulted in the construction of a complete physical map which spans the XLA locus, several advances have been made towards the cloning of this gene. These include:

1) The construction of a physical map of Xq22 which has a coverage of 50-60% of the region, based upon my estimate that Xq22 spans 10-12Mb. This map should therefore help in determining whether any new probes generated will be useful or not in this project since they can easily be incorporated into the existing physical map. Already, the majority of probes used in this study have been show to be telomeric to DXS94 and therefore do not warrant further investigation in terms of XLA eg. for associated polymorphisms for genetic studies.

2) The establishment of a characterised hybrid panel of radiation hybrids which will be a useful resource both for creating and mapping additional DNA markers.

3) The identification of CpG islands close to DXS178 any of which may be associated with candidate genes for XLA.

4) The isolation and mapping of DXS178-positive cosmids as a first step towards cloning the associated islands, and also towards isolating new probes which extend beyond the islands, possibly enabling the completion of the PFGE map of Xq22.
6.2 The Next Steps Towards the Isolation of the XLA Locus

In the previous section (6.1) the main advances presented in this thesis were summarised. I have laid out below the six steps that I would take next in order to build upon the information gained from the present study.

1. The genetic resource available in the XLA families should be maximised. There are several ways to do this. Firstly, the number of XLA families potentially available for linkage studies can be increased by establishing the carrier status of females in families with sporadic XLA patients. This can be done by investigating the X chromosome inactivation patterns in B cells from these individuals.

Fifteen XLA families have been investigated in genetic studies with the markers DXS3, DXS178, DXS94, DXS17 and DXS101. As key females from these families have been uninformative for these markers, I think therefore that we should concentrate on increasing the informativeness of these probes. There is a lower frequency of polymorphic RFLP sites generally on the X chromosome compared to the autosomes (Hofker et al., 1986). At present, all the probes used in XLA studies have two allele systems so are at most only informative in 50% of female carriers of XLA. Currently, the most effective way to improve upon this is to identify highly repetitive \((dC.dA)_n\) sequences close to each of the markers that are tightly linked to the XLA locus. The approach I would chose would be to use the probes to isolate cosmids from the ICRFc100 library used in the present study. These cosmids should then be screened for the presence of \((dC.dA)_n\) repeats which would greatly increase the number of informative heterozygotes for each marker in the XLA families. If possible, more than one \((dC.dA)_n\) sequence could be isolated for each marker. This work has already commenced in the Department with the DXS178-positive cosmids described in Section 5 of this thesis. A \((dC.dA)_n\) repeat has been identified in cosmids.
ICRFc100D1020 and ICRFc100E03167 and is presently being sequenced in order to create PCR primers (Munoree Padayachee, work in progress).

2. *The probe DXS101 should be utilised fully.* There are known to be at least five species of DNA which hybridise to DXS101 within Xq22 (Hofker, et al., 1987) and these are likely to be dispersed within this region judging from the multiple bands (3 or more) obtained in early PFGE results. Cosmids which are positive for DXS101 should therefore be obtained and the different DXS101 species could be distinguished by digesting the cosmids with *Eco*RI, the cloning site enzyme and *Msp*I, the enzyme for which a polymorphism and several constant fragments have been found (see Appendix A1). Single copy probes could then be obtained from the different cosmids which could then be individually localised using PFGE and the radiation hybrid panel described in Section 4.3.

DXS101 has already been excluded from hybridising to the 2.7Mb *Mlu* I band(s) recognised by DXS17, DXS87, DXS94, DXS147 and DXS211, from the 570kb *Mlu* I band detected by PLP and from the 470kb *Bss*H II band to which DXS178 hybridises. Preliminary data has also been obtained by hybridising DXS101 to the hybrid panel used in Section 4.3. The hybridisation of DXS101 to *Msp* I digests indicated that the polymorphic 7.5kb band and also a 2kb constant band were present in hybrid 5 but absent in hybrid 52. This indicated that this particular DXS101 species is located either proximally to DXS94 or distal to DXS83 (see Figure 6.1).

3. *The mapping of the breakpoint in the MBU cell line, in which a new boundary for the XLA locus is suggested, needs to be repeated with adequate controls.* There have been data published indicating that lymphoblastoid lines frequently show alterations in their methylation status (Silva and White, 1988). This work should therefore be repeated with an adequate number of control lymphoblastoid cell lines. A second and simpler possibility would be to attempt
to obtain venous blood directly from MBU and repeat the study. I have recently obtained the probe DXS121, which maps just proximal to the MBU breakpoint; this marker can now be used to corroborate this data.

It would now also be prudent to examine the cell line of TCD patient RvD (see Section 1.4.1.3), in the same way, as the most distal point of this individual’s deletion was estimated by cytogenetic analysis to extend even further into Xq22 than that of MBU. It would therefore be interesting to determine whether or not DXS178 detects altered sized fragments in this cell line since this may reduce the extent of DNA which could contain the XLA locus even further.

4. **The radiation hybrids should be exploited to produce more probes from Xq22.**

Barker *et al.* (1989) have reported that the region Xq22-24 is underrepresented in X chromosome libraries. It would therefore be sensible to make a library from a DNA source containing only the X chromosome region of interest, i.e. the radiation hybrids from Section 4.3.

In order to do this, the hybrids should be scored for the presence of either DXS96 or DXS118, the two markers I have obtained that are deleted in the choroideremia patient MBU but present in choroideremia patient NP and so are distal to DXS3. Hybrids which score positively for one or both of these probes, and also for DXS178 and DXS94 should therefore contain DNA which includes and flanks the XLA locus. Ideally, one or two hybrids which did not score positively for many additional human X chromosome markers (Appendix A3) as a DNA source to produce probes by the use of species-specific Alu primers and the polymerase chain reaction, or by making a cosmid library and screening for Alu-positive clones. Hybrid 129 would be particularly suitable for this as it has the fewest number of positive scores for probes mapping to areas of the X chromosome other than Xq22 (see Appendix A3).
5. The CpG islands close to DXS178 should be cloned. There are two approaches to this: the first is to isolate probes from the ends of the cosmids described in Section 5 and use them to rescreen a cosmid library. The SP6 and T7 promoters within the cosmid vector Lorist X have been used to make riboprobes from cosmid ICRFc100D1020, both of which were found to contain repetitive sequences (M. O'Reilly and A. Sweatman, unpublished observations). In addition, the cosmids have now been investigated thoroughly for the presence of single-copy end fragments, so far without success (Angela Sweatman, work in progress). In view of this a second approach ie. to screen YAC libraries with DXS178 may be preferable. These should firstly contain and span the CpG islands close to DXS178. In addition the entire YAC could be screened against cDNA libraries prepared from both pre-B cell and B-cell lines. I would do this by making cosmids from the YAC inserts, labelling the cosmids and using each to screen against plated-out cDNA libraries. By using the YAC in small segments, the likelihood of detecting small transcripts would be increased.

6. All XLA patients should be screened using PFGE. All patients should be screened using DXS178 in combination with EagI and BssHII. With the exception of CF, all successful attempts to isolate genes using the positional cloning approach have relied upon the identification of one or more patients with gross genetic mutations. In order to maximise the chance of finding a changed band using PFGE, sporadic patients should be investigated in addition to those whose X-linked pedigree has been established.

The T cell stimulation technique has now been established in the department and is effective in producing adequate numbers of T cells (>10⁷) from 1ml blood and from day-old blood samples. This is practical as patients do not have to attend this hospital in order to give blood.

Since the submission of this thesis, I have received information that one DNA species recognised by DXS101 may be within 1200kb of DXS178 (D. Vetrie, Department of Pediatric Genetics, Guy's Hospital, London). This probe
should therefore be used to screen the patients in addition to DXS178.

7. Additional Strategies

I have outlined the steps that I consider a priority to take above, however there are several additional or alternative strategies which could be taken, including exon trapping. This technique makes it possible to screen large stretches of DNA for sequences which are transcribed using an "exon trapping" construct, which detects splice donor sites in tens of kb of DNA at a time (Duyk et al., 1990; Auch and Reth, 1990). Thus it may be viable to screen DXS178 positive YACs or cosmids for transcribed sequences in this way.

An alternative way to increase the saturation of probes in Xq22 would be to make a limited library using a technique such as chromosome band microdissection (Lüdecke et al., 1989). This approach has been used successfully to produce large numbers of clones from the DNA loci such as 11p13 (Wilm's Tumour), 15q11-12 (Prader-Willi syndrome) and from around the fragile X locus in Xq27.3 (Lüdecke et al., 1989; MacKinnon et al., 1990). The clones produced in this manner typically have small inserts and the restriction enzyme used to digest the DNA results in a high proportion of single-copy clones.

6.3 Additional Implications of this Study

In addition to the progress made in refining the XLA locus, the mapping data presented should be of use for linkage or mapping studies involving disease syndromes which currently or are in the future shown to map to this region of the X chromosome. These include cloned genes, such as PLP (Pelizaeus-Merzbacher disease; Gencic et al., 1989), GLA (Fabry disease; Desnick et al., 1987) and COL4α5(IV) (Alport Syndrome; Barker et al., 1990) and disease loci such as X-linked megalocornea which is linked to DXS94 and DXS87 (Chen et al., 1989). The absence of Xq22 male-viable deletions in the literature also
suggests that this chromosome region may be rich in developmentally important genes.

Despite the exponential growth rate in the understanding of B cell lineage maturation, there are few candidate genes for XLA at present. "Positional cloning" therefore continues to represent the most useful approach to isolating the XLA gene.
BIBLIOGRAPHY


Burmeister, M., Monaco, A.P., Gillard, E.F., van Ommen, G-J. B., Affara,


Desnick, R.J., Bernstein, H.S., Astrin, K.H. and Bishop, D.F. (1987) Fabry
Disease: molecular diagnosis of hemizygotes and heterozygotes. Enzyme 38, 54.


Karasuyama, H., Kudo, A. and Melchers, F. (1990) The proteins encoded by the $V_{\text{preB}}$ and lambda5 pre-B cell-specific genes can associate with each other and with $\mu$ heavy chain. J. Exp. Med. 172, 969.


Kudo, A. and Melchers, F. (1987) A second gene, $V_{\text{preB}}$ in the lambda5 locus of the mouse which appears to be selectively expressed in pre-B lymphocytes. EMBO J. 6, 2267.


Nguyen, C., Pontarotti, P., Birnbaum, D., Chimini, G., Rey, J.A., Mattei, J-F. and Jordan, B. (1987) Large scale physical mapping in the q27 region of the human X chromosome: the coagulation factor IX gene and the mcf.2 transforming sequence are separated by at most 270 kilobase pairs and are surrounded by several "HTF islands". EMBO J. 6, 3285.


Schwaber, J., Lazarus, H. and Rosen, F.S. (1978) Restricted classes of


Tsubata, T. and Reth, M. (1990) Products of pre-B cell-specific genes (\(\lambda_5\) and \(\gamma_{mB}\)) and the immunoglobulin \(\mu\) chain form a complex that is transported onto the cell surface. J. Exp. Med. 172, 973.


Appendix A1  Details of probes used in this Study.

<table>
<thead>
<tr>
<th>Probe Name</th>
<th>HGM10 Designation</th>
<th>Insert size/vector/ enzyme</th>
<th>Polymorphic Enzyme</th>
<th>Band size/ Frequency</th>
<th>Washing stringency</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>212/9*</td>
<td>DXS178</td>
<td>10 kb/pUC9/EcoR I</td>
<td>Taq I</td>
<td>3.2 kb 0.7</td>
<td>0.5 x SSC</td>
<td>Cooke et al., 1985</td>
</tr>
<tr>
<td>pXG12</td>
<td>DXS94</td>
<td>1.5 kb/pAT153 Bam HI + EcoR I</td>
<td>Pst I</td>
<td>6.5 kb 0.52</td>
<td>0.5 x SSC</td>
<td>Davatelis et al., 1985</td>
</tr>
<tr>
<td>cX4.4</td>
<td>DXS147</td>
<td>3.6 kb/pAT153/EcoR I ▲3.4 kb Pstl + EcoR I</td>
<td>No</td>
<td>n/a</td>
<td>0.5 x SSC</td>
<td>Hofker et al., 1986</td>
</tr>
<tr>
<td>S21/S9</td>
<td>DXS171</td>
<td>7.2 kb/pBR328/EcoR I</td>
<td>Taq I</td>
<td>2.2 kb 0.65</td>
<td>0.5 x SSC</td>
<td>Drayna et al., 1984</td>
</tr>
<tr>
<td>pA13</td>
<td>DXS87</td>
<td>1.6 kb/pBR328/EcoRI</td>
<td>Bgl II</td>
<td>4.7 kb 3.5 +1.2 kb</td>
<td>0.5 x SSC</td>
<td>ATCC</td>
</tr>
<tr>
<td>G3</td>
<td>DXS88</td>
<td>1.0 kb/pBR328 EcoRI + BamHI</td>
<td>Bgl II</td>
<td>4.7 kb 3.5 kb</td>
<td>0.5 x SSC</td>
<td>ATCC</td>
</tr>
<tr>
<td>pD3</td>
<td>DXS211</td>
<td>1.9 kb/pAT153 Hind III</td>
<td>No</td>
<td>n/a</td>
<td>0.5 x SSC</td>
<td>Cremers et al., 1988</td>
</tr>
<tr>
<td>St3</td>
<td>DXS54</td>
<td>2.1 kb/pBR329/EcoR I</td>
<td>No</td>
<td>n/a</td>
<td>3 x SSC</td>
<td>Oberle et al., '86.</td>
</tr>
</tbody>
</table>
## Appendix A1 continued.

<table>
<thead>
<tr>
<th>Probe Name</th>
<th>HGM10 Designation</th>
<th>Insert size/vector/Enzyme</th>
<th>Polymorphic Enzyme</th>
<th>Band size/Frequency</th>
<th>Washing Stringency</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRL1*</td>
<td>PLP</td>
<td>2.0 kb/pUC9/EcoRI</td>
<td>Msp I</td>
<td>4.5 kb 0.08</td>
<td>1 x SSC</td>
<td>H. Willard, Stanford</td>
</tr>
<tr>
<td>pX5</td>
<td>DXS24</td>
<td>0.8 kb/pAT153/PstI</td>
<td>No</td>
<td>n/a</td>
<td>0.5 x SSC</td>
<td>Balazs et al., 1984</td>
</tr>
<tr>
<td>716</td>
<td>DXS83</td>
<td>0.8 kb/p?/Hind III</td>
<td>No</td>
<td>n/a</td>
<td>1 x SSC</td>
<td>Hofker et al., 1985</td>
</tr>
<tr>
<td>757</td>
<td>DXS114</td>
<td>4 kb/p?/Hind III</td>
<td>No</td>
<td>n/a</td>
<td>0.2 x SSC</td>
<td>Hofker et al., 1985</td>
</tr>
<tr>
<td>cx52.5</td>
<td>DXS101</td>
<td>3.6 kb/pAT153/</td>
<td>Msp1</td>
<td>7.7 kb 0.35</td>
<td>3 x SSC</td>
<td>Hofker et al., 1987</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EcoRI</td>
<td></td>
<td>7.5 kb 0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>212A</td>
<td>(DXS178)</td>
<td>7 kb / Taq I + EcoRI I</td>
<td>No</td>
<td>n/a</td>
<td></td>
<td>This thesis</td>
</tr>
<tr>
<td>212B*</td>
<td>(DXS178)</td>
<td>3 kb / Taq I + EcoRI I</td>
<td>Yes</td>
<td>See DXS178</td>
<td></td>
<td>This thesis</td>
</tr>
<tr>
<td>pH1-14-2*</td>
<td>(COL4A5)</td>
<td>1.3 kb / EcoRI I / pUC18</td>
<td>no</td>
<td>n/a</td>
<td></td>
<td>Karl Tryggvason Biocenter, Oulu</td>
</tr>
</tbody>
</table>

**Key:**

* = cDNA probe

* = Preanneal for 20 mins with sonicated human DNA as described in Materials and Methods.

▲ = the double digest enables separation of insert and vector DNA on a preparative gel.

▼ = washing stringencies are indicated for 30 minute washes at 65°C.

c = constant band

Additional references for the localisations of these probes can be found in Mandel et al., 1989.
### Appendix A2

#### Primary Screening Details for IFGT Hybrids

<table>
<thead>
<tr>
<th>DXS NO.</th>
<th>DXY20</th>
<th>DXS31</th>
<th>STS</th>
<th>DXS278</th>
<th>DMD</th>
<th>DMD</th>
<th>DXS7</th>
<th>DXS146</th>
<th>DXS14</th>
<th>DXS15</th>
<th>DXS17</th>
<th>DXS11</th>
<th>DXS42</th>
<th>DXS37</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAME</td>
<td>29C124</td>
<td>MIA</td>
<td>B12H15</td>
<td>CRI-S232</td>
<td>71.4</td>
<td>J66</td>
<td>L1.28</td>
<td>pTAK8B</td>
<td>P58-1</td>
<td>47Z</td>
<td>S21</td>
<td>P22-33</td>
<td>P43-15</td>
<td>3OR1b</td>
</tr>
<tr>
<td>LOCATION</td>
<td>p22.32</td>
<td>p22.32</td>
<td>p22.32</td>
<td>p22.32</td>
<td>p21.3-21.2</td>
<td>p11.4-11.3</td>
<td>11.22</td>
<td>p11.22</td>
<td>q21</td>
<td>q22</td>
<td>q24-q25</td>
<td>q24-25</td>
<td>q25-26</td>
<td></td>
</tr>
<tr>
<td>HYBRID I.D. NO.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>nd</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>52</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>92</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>nd</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>129</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>143</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>203</td>
<td>1</td>
<td>nd</td>
<td>0</td>
<td>nd</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DXS NO.</th>
<th>HPRT</th>
<th>DXS144E</th>
<th>DXS51</th>
<th>F9</th>
<th>DXS105</th>
<th>DXS98</th>
<th>DXS90</th>
<th>DXS304</th>
<th>DXS305</th>
<th>DXS52</th>
<th>DXS15</th>
<th>35.516</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAME</td>
<td>PHPT30</td>
<td>c11</td>
<td>S2A</td>
<td>FIX-P1</td>
<td>cX55.7</td>
<td>4Q-8</td>
<td>pXG-20</td>
<td>U6.2</td>
<td>S135.691</td>
<td>S14-1</td>
<td>DX13</td>
<td>F6C-G6PG</td>
</tr>
<tr>
<td>LOCATION</td>
<td>q26</td>
<td>q26</td>
<td>q26</td>
<td>q26</td>
<td>q26.2-26.3</td>
<td>q26.3-27.1</td>
<td>q27.1-27.2</td>
<td>q27.2</td>
<td>q27-28</td>
<td>q28</td>
<td>q28</td>
<td>q28</td>
</tr>
<tr>
<td>HYBRID I.D. NO.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>nd</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>nd</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>nd</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>52</td>
<td>nd</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>92</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>129</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>143</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>203</td>
<td>nd</td>
<td>0</td>
<td>nd</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>nd</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Loci for probes were updated according to Mandel et al., 1989.

Key: 1 = positive score
2 = weak positive score
nd = not done
Appendix 3
Restriction map of Lorist 4

There are no restriction sites for SalI, KpnI, MluI, SfiI or NotI.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>No. of sites</th>
<th>Position of Site in vector</th>
<th>Size of fragment</th>
</tr>
</thead>
<tbody>
<tr>
<td>BamHI</td>
<td>1</td>
<td>624</td>
<td>5385</td>
</tr>
<tr>
<td>BssHII</td>
<td>1</td>
<td>4756</td>
<td>4132, 1253</td>
</tr>
<tr>
<td>ClaI</td>
<td>1</td>
<td>2</td>
<td>623, 1253</td>
</tr>
<tr>
<td>EcoRI</td>
<td>3</td>
<td>529, 918, 2162</td>
<td>294, 1244, 3752, 95</td>
</tr>
<tr>
<td>HindIII</td>
<td>1</td>
<td>617</td>
<td>5378, 7</td>
</tr>
<tr>
<td>Nael</td>
<td>2</td>
<td>4861, 5144</td>
<td>283, 865, 4237</td>
</tr>
<tr>
<td>NarI</td>
<td>2</td>
<td>4111, 4359</td>
<td>3487, 248, 1650</td>
</tr>
<tr>
<td>SacII</td>
<td>1</td>
<td>246</td>
<td>378, 5007</td>
</tr>
</tbody>
</table>

❖ indicates the BamHI cloning site. This table can be used to determine the orientation of the fragments obtained by enzyme digestion of cosmids D1020 and E03167. For example, with a BssHII digest, 4123 bp of Lorist 4 will be linked to one insert fragment and 1253 bp to another. The relative intensities of the signals obtained upon hybridisation with Lorist will differ by 3-4 fold.