Molecular Genetics of Chorea-Acanthocytosis

Carol N. M. Dobson-Stone

Wolfson College
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy
Hilary 2004
Chorea-acanthocytosis (ChAc) is an autosomal recessive neurological disorder whose characteristic features include hyperkinetic movements and abnormal red blood cell morphology. The disorder shares features with Huntington’s disease and McLeod syndrome (MLS), and can sometimes be difficult to distinguish clinically from the latter. In 1997, ChAc was linked to a 6-cM region on chromosome 9q21-22. A novel gene, \textit{CHAC}, was identified in the critical region. \textit{CHAC} (now renamed \textit{VPS13A}) encodes a large protein called chorein, with a yeast homologue implicated in protein sorting.

In this study, all 73 exons plus flanking intronic sequence in \textit{VPS13A} were screened for mutations in 83 unrelated ChAc patients. We identified 88 different \textit{VPS13A} mutations in 72 probands, comprising six deletions of entire exons, 22 nonsense, 36 frameshift, 19 splice-site and five missense mutations. This disorder therefore shows substantial allelic heterogeneity: however, evidence for common inheritance of the EX70_73del mutation in four French Canadian pedigrees indicates a possible founder effect in this population.

Expression of \textit{VPS13A} appears to be ubiquitous, as determined by tissue-specific analysis of mRNA and chorein distribution. However, chorein expression was markedly reduced or undetectable in lymphoblasts, fibroblasts and erythrocyte membranes from 14 ChAc patients. In contrast, MLS cells showed chorein expression similar to control levels, suggesting that loss of chorein expression is a diagnostic feature of ChAc.

Yeast two-hybrid analysis of six different \textsim{sim}600 amino-acid chorein fragments was used to screen a human brain cDNA library for proteins that may interact with chorein. One fragment interacted weakly with constructs derived from transcription factor NF-κB, putative protein phosphatase PP2C\eta and TAB2, a protein implicated in the mitogen-activated kinase cascade. Although exogenously expressed chorein and TAB2 did not appear to colocalise, co-immunoprecipitation experiments supported an interaction between the two proteins, suggesting an avenue for future research into chorein function.

\textit{Submitted for the degree of Doctor of Philosophy, Hilary 2004}
Declaration

All the work in this thesis was performed by the author in the Wellcome Trust Centre for Human Genetics, Faculty of Clinical Medicine, University of Oxford, except as detailed below:

**Mutational analysis of VPS13A in patients with chorea-acanthocytosis (Chapter 3), Wave I:** Luca Rampoldi was responsible for cloning the \textit{CHAC} gene (\textit{VPS13A}). I assisted him with mutation detection of the original 11 ChAc probands. I also reconfirmed the mutations in these probands by restriction digest or sequencing, and checked that mutations co-segregated with the disease haplotypes in the respective families.

**Wave II:** Lorne Lonie performed the DHPLC analysis of the Wave II cohort. I performed PCR amplification and hybridisation of the proband samples, and all post-DHPLC procedures.

**Wave III:** Lorne Lonie performed the DHPLC analysis of the Wave III cohort. Under my supervision, Lea Filippone and Sarah Westbury performed PCR amplification and hybridisation of the proband samples, and sequencing analysis of DHPLC variants. I reconfirmed the mutations in these probands by sequencing, and checked that mutations co-segregated with the disease haplotypes in the respective families.

The work reported in this thesis has not been submitted for any other degree in this or any other university or institute of learning.
Acknowledgements

During the course of this study, I was the recipient of a Wellcome Trust Prize Studentship. I would like to thank the patients and their families, who have made this study possible, and the numerous clinicians who were involved in patient sample collection. In particular, Prof Adrian Danek was the driving force behind collection of patient samples and clinical data, and was always ready to give help and advice on the clinical aspects of this study.

I would also like to thank Prof Tony Monaco for his brilliant supervision and encouragement throughout my time at the Wellcome Trust Centre. Despite the many demands on his time, he was always quick to reply with sensible answers to my stupid questions, and has provided the gold standard by which I shall measure all future bosses! Drs Luca Rampoldi and Antonio Velayos-Baeza, the former and current post-doc on the ChAc project, have also been extremely generous with their time and wisdom and have been instrumental in the success of this study.

Additional thanks go to the rest of the Monaco group for all their help in scientific and non-scientific matters, and for making the past six years at the Wellcome Trust Centre seem more like fun than work. Our lab neighbours, the Genomics group and the former Hovnanian group, were always willing to give technical advice and much-needed reagents. Outside the lab, I am indebted to Jan Trinder for all her help and the ‘lucky touch’ that ensured acceptance of my manuscripts! My non-scientific friends have always been supportive, and listened politely without yawning to my moans about experiments not going to plan, for which I am very grateful.

Above all, I would like to thank my parents and the rest of my rapidly expanding family for all their love, and the support they have shown in my every endeavour since I was old enough to hold a pen. Special thanks to my husband, without whom I wouldn’t have eaten or otherwise functioned properly over the past three months, and for always being there to rescue me from irretrievable geekiness.
Glossary of abbreviations used in this thesis

5FOA 5-fluoroorotic acid
aa amino acid
ABL abetalipoproteinemia
AD activation domain
AT annealing temperature
BCA bicinchonic acid
BLAST basic local alignment sequence tool
bp base pair
BSA bovine serum albumin
C- carboxy
cDNA complementary DNA
CGD chronic granulomatous disease
ChAc chorea-acanthocytosis
CK creatine kinase
cM centimorgan
CMV cytomegalovirus
CT computed tomography
DB DNA-binding domain
DHPLC denaturing high-performance liquid chromatography
dH₂O deionised water
DMD Duchenne muscular dystrophy
DNA deoxyribonucleic acid
dNTP deoxynucleoside-5’-triphosphate
DTT dithiothreitol
EBV Epstein-Barr virus
EDTA ethylenediaminetetraacetic acid
EGFP enhanced green fluorescent protein
EST expressed sequence tag
gDNA genomic DNA
GST glutathione-S-transferase
h hour
HARP hypoprebetalipoproteinemia, acanthocytosis and retinitis pigmentosa
HBL familial hypobetalipoproteinemia
HD Huntington’s disease
HDL2 Huntington’s disease-like 2
-His histidine-deficient
HSS Hallervorden-Spatz syndrome
IKK 1κB kinase
IκB inhibitor of NF-κB
IL-1 interleukin-1
IPTG isopropylthio-β-D-galactoside
kb kilobase pair
kDa kilodalton
LB Luria broth
-Leu leucine-deficient
LiAc lithium acetate
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>mA</td>
<td>milliamps</td>
</tr>
<tr>
<td>MAP</td>
<td>mitogen-activated protein</td>
</tr>
<tr>
<td>MAPKKK</td>
<td>MAP kinase kinase kinase</td>
</tr>
<tr>
<td>mg</td>
<td>milligram</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>ml</td>
<td>millilitre</td>
</tr>
<tr>
<td>MLS</td>
<td>McLeod syndrome</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>MTP</td>
<td>microsomal triglyceride transfer protein</td>
</tr>
<tr>
<td>Mw</td>
<td>molecular weight</td>
</tr>
<tr>
<td>μg</td>
<td>microgram</td>
</tr>
<tr>
<td>μl</td>
<td>microlitre</td>
</tr>
<tr>
<td>μm</td>
<td>micrometer (micron)</td>
</tr>
<tr>
<td>μM</td>
<td>micromolar</td>
</tr>
<tr>
<td>N-</td>
<td>amino</td>
</tr>
<tr>
<td>NF-kB</td>
<td>nuclear transcription factor kappa B</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>nm</td>
<td>nanometre</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>OPMD</td>
<td>oculopharyngeal muscular dystrophy</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PDI</td>
<td>protein disulfide isomerase</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PET</td>
<td>positron emission tomography</td>
</tr>
<tr>
<td>PKAN</td>
<td>pantothenate kinase-associated neurodegeneration</td>
</tr>
<tr>
<td>pmol</td>
<td>picomole</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethylsulphonyl fluoride</td>
</tr>
<tr>
<td>PP2Cη</td>
<td>protein phosphatase 2C eta</td>
</tr>
<tr>
<td>PTC</td>
<td>premature termination codon</td>
</tr>
<tr>
<td>PVC</td>
<td>prevacuolar compartment</td>
</tr>
<tr>
<td>RFLP</td>
<td>restriction fragment length polymorphism</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNAse</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcriptase polymerase chain reaction</td>
</tr>
<tr>
<td>s</td>
<td>second</td>
</tr>
<tr>
<td>sarkosyl</td>
<td>N-lauryl sarcosine, sodium salt</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulphate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>STE</td>
<td>saline Tris-EDTA</td>
</tr>
<tr>
<td>SV40</td>
<td>simian virus 40</td>
</tr>
<tr>
<td>TAB2</td>
<td>TAK1-binding protein 2</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-acetate EDTA</td>
</tr>
<tr>
<td>TAK1</td>
<td>transforming growth factor β-activated kinase 1</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-borate EDTA</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris-buffered saline</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>TEMED</td>
<td>tetramethylethylenediamine</td>
</tr>
<tr>
<td>TF</td>
<td>transcription factor</td>
</tr>
<tr>
<td>TGN</td>
<td>trans-Golgi network</td>
</tr>
<tr>
<td>TPR</td>
<td>tetratricopeptide repeat</td>
</tr>
<tr>
<td>-Trp</td>
<td>tryptophan-deficient</td>
</tr>
<tr>
<td>U</td>
<td>unit</td>
</tr>
<tr>
<td>UBA</td>
<td>ubiquitin-associated</td>
</tr>
<tr>
<td>-Ura</td>
<td>uracil-deficient</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>vps</td>
<td>vacuolar protein sorting</td>
</tr>
<tr>
<td>X-Gal</td>
<td>5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside</td>
</tr>
<tr>
<td>Y2H</td>
<td>yeast two-hybrid</td>
</tr>
</tbody>
</table>
# Contents

Abstract .......................................................................................................................... i
Declaration ....................................................................................................................... ii
Acknowledgements ......................................................................................................... iii
Glossary of abbreviations used in this thesis ................................................................. iv
Contents .......................................................................................................................... vii
List of tables and figures ............................................................................................... xiii

## Chapter 1: Introduction ............................................................................................... 1

1.1 Clinical features of chorea-acanthocytosis ............................................................... 1

1.1.1 General features of chorea-acanthocytosis .......................................................... 2
1.1.2 Movement disorders in ChAc ................................................................................. 3
1.1.3 Neuropsychiatric features ..................................................................................... 4
1.1.4 Brain pathology ...................................................................................................... 4
1.1.5 Peripheral nerve and muscle pathology ................................................................. 6
1.1.6 Laboratory findings ............................................................................................... 6
1.1.7 Treatment ............................................................................................................. 7

1.2 Other neuroacanthocytoses: McLeod syndrome ....................................................... 8

1.2.1 Clinical features ................................................................................................... 9
1.2.2 Genetics ............................................................................................................... 10
1.2.3 Possible functions of XK ..................................................................................... 11

1.3 Other neuroacanthocytoses: abetalipoproteinemia .................................................... 14

1.3.1 Clinical features .................................................................................................. 14
1.3.2 Treatment ........................................................................................................... 15
1.3.3 Genetics ............................................................................................................. 16
1.3.4 Function of MTP ............................................................................................... 16

1.4 Other neuroacanthocytoses: hypobetalipoproteinemia ............................................. 18

1.4.1 Genetics ............................................................................................................. 18

1.5 Other syndromes associated with acanthocytosis ..................................................... 19

1.5.1 Hallervorden-Spatz syndrome .......................................................................... 19
1.5.2 Huntington’s disease-like 2 .............................................................................. 20

1.6 The basis of acanthocytosis ..................................................................................... 21
1.6.1 Erythrocyte structure ................................................................. 21
1.6.2 Normal and aberrant erythrocyte morphology............................ 22
1.6.3 Detection of acanthocytes ............................................................ 22
1.6.4 Factors in acanthocyte formation ............................................... 23
1.6.5 Role of lipids ............................................................................. 23
1.6.6 Role of proteins .......................................................................... 25
1.7 Genetics of ChAc ......................................................................... 27
1.7.1 Mode of inheritance .................................................................... 27
1.7.2 CHAC positional cloning ............................................................. 27
1.7.3 Genomic structure of VPS13A ..................................................... 28
1.8 Chorein ....................................................................................... 29
1.8.1 Chorein orthologues ................................................................... 30
1.8.2 The human VPS13 family ........................................................... 32
1.9 Research strategy ......................................................................... 34

Chapter 2: Materials and Methods .................................................. 35

2.1-2.10 NUCLEIC ACID TECHNIQUES ............................................. 35
2.1 Preparation of nucleic acids ......................................................... 35
2.1.1 Extraction of genomic DNA from blood ....................................... 35
2.1.2 Extraction of gDNA from lymphoblastoid/fibroblast cell lines .......... 36
2.1.3 Preparation of recombinant DNA from bacteria - minipreps .......... 36
2.1.4 Preparation of recombinant DNA from bacteria - maxipreps .......... 36
2.1.5 Extraction of RNA ..................................................................... 37
2.2 Polymerase chain reaction (PCR) .................................................. 37
2.2.1 Standard PCR ........................................................................... 37
2.2.2 PCRs for denaturing high-performance liquid chromatography
analysis ........................................................................................................ 38
2.2.3 Long template PCR ................................................................... 38
2.2.4 Colony PCR ............................................................................... 39
2.2.5 Reverse-transcription PCR (RT-PCR) ......................................... 39
2.3 Agarose gel electrophoresis .......................................................... 39
2.4 Quantification of DNA and RNA .................................................... 40
2.4.1 Spectrophotometric quantification .......................................................... 40
2.4.2 Quantification on agarose gel ................................................................. 41

2.5 PCR product/restriction digest purification ............................................... 41

2.6 Nucleic acid hybridisation protocols ........................................................ 42
  2.6.1 Southern blot detection ....................................................................... 42
  2.6.2 Northern blot detection ....................................................................... 43

2.7 DNA sequencing analysis ......................................................................... 44
  2.7.1 Sequencing on ABI 377 ...................................................................... 44
  2.7.2 Sequencing on ABI 3700 .................................................................... 45

2.8 Mutation detection protocols ................................................................... 46
  2.8.1 Strategy ............................................................................................... 46
  2.8.2 DHPLC analysis ................................................................................ 46

2.9 Haplotype analysis .................................................................................. 47

2.10 Molecular cloning .................................................................................... 48
  2.10.1 Restriction digestion ......................................................................... 48
  2.10.2 Modification of restriction fragment ends ......................................... 49
  2.10.3 Ligation ............................................................................................. 49
  2.10.4 Transformation of E. coli ................................................................. 50
  2.10.5 Preparation and revival of glycerol stocks ...................................... 52

2.11 CELL CULTURE TECHNIQUES .................................................................. 54
  2.11.1 Cell lines .......................................................................................... 54
  2.11.2 Maintenance of cells ........................................................................ 54
  2.11.3 Harvesting and subculture of adherent cells ...................................... 55
  2.11.4 Harvesting and subculture of non-adherent cells .............................. 55
  2.11.5 Preparation and revival of frozen stocks ............................................. 55
  2.11.6 Transfection of 293T and MRC5-SV2 cells ..................................... 56

2.12-2.20 PROTEIN TECHNIQUES ................................................................ 57
  2.12 Preparation of recombinant antigen ..................................................... 57
    2.12.1 Expression of recombinant protein in bacterial cultures ............... 57
    2.12.2 Fractionation of bacterial cell lysates .......................................... 57
    2.12.3 Solubilisation of glutathione-S-transferase (GST)-chorl .............. 58
    2.12.4 Purification of GST-chorl .............................................................. 58
    2.12.5 Antigen concentration and immunisation ..................................... 60
  2.13 Protein sample preparation ................................................................. 61
2.13.1 Yeast cell lysate preparation .............................................................. 61
2.13.2 Mammalian cell lysate preparation .................................................... 61
2.13.3 Erythrocyte membrane preparation ...................................................... 62

2.14 Dot blotting ............................................................................................ 63

2.15 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) .................................... 63
   2.15.1 Standard PAGE .................................................................................. 63
   2.15.2 Gradient PAGE .................................................................................. 64

2.16 Coomassie staining .................................................................................. 64

2.17 Western blot analysis .............................................................................. 64
   2.17.1 Transfer of proteins ........................................................................... 64
   2.17.2 Immunodetection ............................................................................. 65

2.18 Antibody purification and processing ....................................................... 66
   2.18.1 Column preparation .......................................................................... 66
   2.18.2 Serum treatment ............................................................................. 66
   2.18.3 Antibody depletion assay ................................................................... 67

2.19 Co-immunoprecipitation ........................................................................ 68

2.20 Immunofluorescence (IF) ........................................................................ 69

2.21 YEAST TWO-HYBRID PROTOCOLS ....................................................... 72
   2.21.1 Y2H library screen strategy .............................................................. 72
   2.21.2 Retransformation strategy .................................................................. 73
   2.21.3 Assembly of Y2H bait and prey constructs ....................................... 73
   2.21.4 Preparation of Y2H cDNA library .................................................... 74
   2.21.5 Small-scale polyethylene glycol/lithium acetate (PEG/LiAc) transformation of yeast .......................................................... 74
   2.21.6 Large-scale PEG/LiAc transformation of yeast .................................. 75
   2.21.7 Extraction recombinant DNA from yeast ........................................... 76

Chapter 3: Mutational analysis of VPS13A in patients with chorea-acanthocytosis .......................................................... 78

3.1 Wave I .................................................................................................... 78
   3.1.1 Description of mutations ................................................................... 78

3.2 Wave II .................................................................................................... 81
5.1.1 Construction of yeast two-hybrid baits .................................................. 105
5.1.2 Investigation of potential homotypic interactions ............................. 106
5.1.3 Yeast two-hybrid library screen ............................................................. 107
5.1.4 Characterisation of tertiary positives .................................................. 109
5.1.5 Further Y2H investigation of chacY2 - PP2Cγ interaction ................. 109
5.1.6 Further Y2H investigation of chacY2 - NF-κB interaction ................. 111
5.2 Investigation of chorein interaction with TAB2 ....................................... 111
5.2.1 Co-immunoprecipitation of chorein and TAB2 ................................. 112
5.2.2 Immunolocalisation of chorein-FLAG and T7-TAB2 ....................... 113
5.3 Summary ................................................................................................. 114

Chapter 6: Discussion ........................................................................ 115
6.1 Mutational analysis of VPS13A ................................................................. 115
6.1.1 VPS13A is the gene mutated in chorea-acanthocytosis ....................... 115
6.1.2 Mode of inheritance of ChAc ............................................................... 116
6.1.3 VPS13A mutation distribution ............................................................ 118
6.1.4 Recurrent VPS13A mutations ............................................................. 119
6.1.5 Predicted effects of VPS13A mutations on function ......................... 124
6.2 Analysis of VPS13A expression ............................................................... 127
6.2.1 Distribution of VPS13A mRNA and protein ....................................... 127
6.2.2 Expression of mutant chorein in ChAc cell lines ................................. 127
6.2.3 Chorein detection for diagnosis of ChAc ............................................. 128
6.3 Identification of potential chorein-interacting proteins ......................... 130
6.3.1 Limitations of the assay ..................................................................... 130
6.3.2 Potential chorein-binding partners .................................................... 131
6.3.3 Possible relationships between chorein-interacting proteins ............. 132
6.4 Potential functions for chorein ............................................................... 135
6.4.1 Chorein: a protein sorter? ................................................................. 135
6.4.2 Possible redundancy in function of VPS13 proteins ....................... 137
6.4.3 Future analysis of chorein function .................................................... 137

References ................................................................................................. 139

Appendix 1: PCR primers used in this study ............................................ 159
A1.1 Primers for amplification of VPS13A exons ....................................... 159
A1.2 Primers for amplification of VPS13A cDNA ..................................... 161
List of tables and figures

Index of Tables

Table 2.1 - Composition of reagents used for nucleic acid techniques............................ 53
Table 2.2 - Composition of reagents used for protein techniques ...................................70
Table 2.3 - Composition of reagents used for yeast two-hybrid techniques ....................77
Table 3.1 - VPS13A disease mutations identified in Wave I...........................................80
Table 3.2 - VPS13A disease mutations identified in Wave II.......................................... 85
Table 3.3 - VPS13A disease mutations identified in Wave III ........................................90
Table 3.4 - Exonic polymorphisms identified in VPS13A..............................................92
Table 5.1 - Number of clones processed in chorein yeast two-hybrid screens .............. 108
Table 6.1 - VPS13A mutations identified in two or more ChAc probands .................... 123

List of Figures

Figure 1.1 - Clinical features of neuroacanthocytoses
Figure 1.2 - Normal and abnormal erythrocyte structure
Figure 1.3 - Physical map of the CHAC region on chromosome 9q21-q22 and
     genomic structure of VPS13A
Figure 1.4 - Putative chorein orthologues
Figure 1.5 - The human VPS13 family

Figure 3.1 - Identification of \textit{VPS13A} mutations in patients with ChAc
Figure 3.2 - Analysis of \textit{VPS13A} mutations in pseudodominant ChAc families
Figure 3.3 - Analysis of gross deletions in \textit{VPS13A}
Figure 3.4 - Further characterisation of \textit{VPS13A} EX70_73del mutation

Figure 4.1 - Expression analysis of \textit{VPS13A} mRNA
Figure 4.2 - Expression of GST-fusion proteins in bacteria
Figure 4.3 - Expression and purification of GST-chor1
Figure 4.4 - Dot blot analysis of anti-chor1 antiserum
Figure 4.5 - Detection of exogenous and endogenous chorein
Figure 4.6 - Purification of anti-chor1 antiserum
Figure 4.7 - Expression of chorein in different cell types
Figure 4.8 - Analysis of chorein expression in ChAc cell lines
Figure 4.9 - Analysis of chorein expression in red blood cells

Figure 5.1 - Construction and expression of baits for yeast two-hybrid analysis
Figure 5.2 - Construction of prey vectors and investigation of homotypic interactions in chorein
Figure 5.3 - Retransformation assays of secondary positives from chacY2 yeast two-hybrid screen
Figure 5.4 - Investigation of the interaction between chacY2 and PP2C\eta
Figure 5.5 - Investigation of the interaction between chacY2 and NF-\kappaB
Figure 5.6 - Co-immunoprecipitation of T7-TAB2 and chac2-myc
Figure 5.7 - Co-immunoprecipitation of T7-TAB2 and chorein-FLAG
Figure 5.8 - Subcellular localisation of TAB2 and chorein

Figure 6.1 - Spectrum of \textit{VPS13A} mutations identified in this study
Figure 6.2 - Haplotype analysis of French Canadian ChAc families
Figure 6.3 - Alignment of \textit{VPS13A} missense mutations
Figure 6.4 - Features of human VPS13 proteins
Chapter 1: Introduction

‘Acanthocyte’, meaning ‘thorny cell’, is a descriptive term for a particular type of abnormal erythrocyte characterised by irregular cytoplasmic projections. Secondary acanthocytosis can occur post-splenectomy (Ward 1979), and as a result of malnutrition (due to, e.g., anorexia) (Takeshita et al. 2002), liver disease (Chitale et al. 1998), or myelodysplasia (Doll et al. 1989). Hereditary acanthocytosis is found together with neurological abnormalities in a rare constellation of disorders, collectively termed neuroacanthocytosis. Chorea-acanthocytosis, the subject of this thesis, is one of these disorders. Other neuroacanthocytoses include McLeod syndrome, abetalipoproteinemia and homozygous hypobetalipoproteinemia. Additional diseases, such as Hallervorden-Spatz syndrome and Huntington’s disease-like 2, can occasionally include acanthocytosis as a phenotype.

1.1 Clinical features of chorea-acanthocytosis

In 1968, a North American family was described with circulating acanthocytes and a distinctive neurological disorder (Estes et al. 1967). Such an association had previously been observed in the recessive disorder abetalipoproteinemia (Bassen and Kornzweig 1950; Schwartz et al. 1961). However, no abnormality in serum lipids was detected in this family, and in another pedigree with similar symptoms: reports on these two families in 1968 concluded that they were suffering from a clinically distinct syndrome (Critchley et al. 1968; Levine et al. 1968). This disorder has been termed variously
Levine-Critchley syndrome (Sakai et al. 1985), amyotrophic chorea with acanthocytosis (Kito et al. 1980), neuroacanthocytosis (Yamamoto et al. 1982) and choreoacanthocytosis (Sakai et al. 1981), but it is now most commonly known as chorea-acanthocytosis.

1.1.1 General features of chorea-acanthocytosis

Chorea-acanthocytosis (ChAc, OMIM 200150) is a multisystem degenerative neurological disorder, characterised by progressive dyskinesia and aberrant erythrocyte morphology. No estimates of its prevalence have been made, but the disorder appears to be more common in Japan (Kuroiwa et al. 1984; Oshima et al. 1985): approximately 50 of the 200 or so cases reported to date are Japanese. ChAc does not appear to be geographically restricted: cases have been reported in Northern and Southern Europe (Sotaniemi 1983; Serra et al. 1986), North and South America (Critchley et al. 1968; Gross et al. 1985), India (Bharucha and Bharucha 1989), Africa (Feinberg et al. 1991), China (Ong et al. 1989) and the Middle East (Bohlega et al. 2003). The onset of neurological symptoms is usually delayed until 25-45 years of age, but can occur as early as the first decade or as late as the seventh (Hardie et al. 1991; Rampoldi et al. 2002). The course of the disease is progressive and death usually occurs after 7-14 years, often due to respiratory complications directly or indirectly linked to the movement disorder. It is not rare for patients to commit suicide (Alonso et al. 1989; Johnson et al. 1998). There is as yet no long-term treatment for this disorder.
1.1.2 Movement disorders in ChAc

As the name implies, chorea is often the dominant neurological symptom in ChAc. Chorea is commonly described as random, abrupt, irregular, involuntary movements that flow from body part to body part in a chaotic manner. Chorea often resembles fragments of normal movement, and many individuals with chorea incorporate the involuntary movements into motor patterns that appear voluntary. Choreatic limb movements in ChAc patients broadly resemble those seen in Huntington’s disease (HD) (Shibasaki et al. 1982). In addition, the orofacial region is usually affected; in contrast with HD, self-mutilation of tongue and lips (Sakai et al. 1981; Kuroiwa et al. 1984) is often seen. Orofacial dyskinesia can also lead to dysarthria and dysphagia, commonly in the form of feeding dystonia. Vocalisations and motor tics reminiscent of those seen in Tourette syndrome are often observed (Hardie et al. 1991); however, coprolalia is reported in only a few cases (Bird et al. 1978; Spitz et al. 1985).

Dystonia, characterised by involuntary muscle spasms that produce twisting postures of different body parts, is also common: it has been reported as the presenting symptom in a few cases (de Yebenes et al. 1988; Peppard et al. 1990; Hardie et al. 1991). The hyperkinetic state can gradually progress to parkinsonism (Sakai et al. 1985; Hardie et al. 1991; Ishikawa et al. 2000) - this has been explained as an initial supersensitivity of brain dopamine receptors followed by progressive pre- and postsynaptic degeneration (Spitz et al. 1985). However, rare patients presenting with parkinsonism have also been reported (Spitz et al. 1985; Bostantjopoulou et al. 2000). In short, the gamut of involuntary movements associated with basal ganglia dysfunction is exhibited in ChAc.
1.1.3 Neuropsychiatric features

Epileptic seizures are common in ChAc - a study of 40 Japanese cases reported seizures as the second most common presenting symptom after chorea (Kuroiwa et al. 1984), and they have been the presenting symptom in European cases (Malandrini et al. 1993; Aasly et al. 1999). A neuropsychiatric review of nine ChAc patients revealed that personality change, with impulsive and distractible behaviour or apathy was most common, and obsessive-compulsive features might also occur (Kartsounis and Hardie 1996). Other psychiatric and behavioural features seen in ChAc include aggression (Alonso et al. 1989; Malandrini et al. 1993), anxiety (Feinberg et al. 1991), depression (Rinne et al. 1994b), disinhibition (Sorrentino et al. 1999) and lying (Yamamoto et al. 1982). Cognitive impairment is common, with deficits in memory, attention and executive skills observed (Kartsounis and Hardie 1996), but intellectual deterioration appears to be mild compared with HD (Medalia et al. 1989).

1.1.4 Brain pathology

As expected from the dyskinesia associated with the disorder, ChAc appears to be primarily a disease of the basal ganglia. Typical post-mortem analyses reveal marked atrophy and astrogliosis of the caudate and putamen, with associated enlargement of the lateral ventricles, similar findings to that in HD. The globus pallidus is also often atrophied, but to a lesser extent (Bird et al. 1978; Sobue et al. 1986; Alonso et al. 1989; Feinberg et al. 1991). Rinne et al. (1994) observed that the thalamus, substantia nigra and anterior horns of the spinal cord were abnormal in 30-50% of previous brain autopsies; and that the subthalamic nucleus, midbrain, pons, medulla, cerebellum and cerebral cortex were apparently unaffected. However, ocular motor abnormalities suggest that the brainstem is involved in the pathology of some ChAc cases.
(Gradstein et al. 2003). More extensive nigral degeneration was observed in post-mortem studies of two cases with overt parkinsonian features (Rinne et al. 1994a).

Computed tomography (CT) and magnetic resonance imaging (MRI) studies are consistent with the post-mortem analyses (see Figure 1.1a). MRI shows increased signal in caudate and putamen on T2-weighted images (Hardie et al. 1991; Malandrini et al. 1993; Tanaka et al. 1998). Serra et al. (1987) used CT to measure caudate size in five patients with suspected ChAc and found that the extent of shrinkage was very similar to that in HD patients. Mild diffuse cerebral atrophy has been observed using CT (Sotaniemi 1983; Spitz et al. 1985; Malandrini et al. 1993; Hirayama et al. 1997), but the degree of cerebral cortex degeneration observed in HD (Vonsattel et al. 1985) does not appear to be typical for ChAc.

Positron emission tomography (PET) findings show marked glucose hypometabolism of caudate and putamen in ChAc patients (Hosokawa et al. 1987; Dubinsky et al. 1989). Tanaka et al. (1998) found global reduction in cerebral blood flow and oxygen metabolism, with marked reduction in caudate and putamen, and less in frontal lobes. Another study showed significant blood flow decrease in the frontal lobe only (Delecluse et al. 1991). An $^{18}$F-fluorodopa-uptake PET study showed evidence of selective degeneration of dopaminergic projections from the nigra compacta to the posterior putamen in neuroacanthocytosis patients; altered $^{11}$C-raclopride uptake implied loss of striatal dopamine D2 receptors (Brooks et al. 1991). The authors suggested that this could explain why choreic and parkinsonian signs are both seen in ChAc.
1.1.5 *Peripheral nerve and muscle pathology*

Reduction or absence of deep tendon reflexes, indicative of a motor neuropathy and not observed in HD, is highly characteristic of ChAc (Sakai *et al.* 1981; Sotaniemi 1983). There is negligible reduction of motor and sensory nerve conduction velocity (Sobue *et al.* 1986; Vita *et al.* 1989; Tanaka *et al.* 1998). Nerve biopsies show a moderate decrease in density of myelinated fibres (Ohnishi *et al.* 1981; Vita *et al.* 1989). Muscle atrophy is a common feature of ChAc and can sometimes occur before dyskinesia (Kito *et al.* 1980; Serra *et al.* 1986). Muscle CT scans of two ChAc patients showed selective symmetrical lower leg atrophy (Ishikawa *et al.* 2000). Electromyograms of suspected ChAc patients generally show high amplitude and long duration potentials, indicating neurogenic abnormalities (e.g., (Ohnishi *et al.* 1981; Gross *et al.* 1985; Vita *et al.* 1989; Tanaka *et al.* 1998)). Muscle biopsies are also compatible with a chronic neurogenic muscular atrophy (Ohnishi *et al.* 1981; Serra *et al.* 1986; Vita *et al.* 1989). When myopathic changes are observed in ChAc, they are considered secondary to the chronic denervation (Limos *et al.* 1982).

1.1.6 *Laboratory findings*

Apart from acanthocytosis, the most common laboratory findings in ChAc are elevated serum muscle-specific creatine kinase (CK) levels (Danek *et al.* 2004). Levels of other enzymes, such as muscle and liver transaminases and lactate dehydrogenase are sometimes also elevated (Serra *et al.* 1986; Tanaka *et al.* 1998). These findings may be due to an often subclinical muscular atrophy (Danek *et al.* 2004). Two ChAc patients showed increased levels of tissue transglutaminase, a cross-linking enzyme involved in assembly of macromolecular structures (Melone *et al.* 2004).
al. 2002). One female with confirmed ChAc showed absence of serum aprebetalipoprotein (Bohlega et al. 1998; Bohlega et al. 2003).

1.1.7 Treatment

Medication for ChAc is aimed primarily at relieving dyskinesia and psychiatric symptoms, or controlling epileptic seizures. Although the latter is often successful, standard neuroleptics are often of little effect (Wihl et al. 2001; Burbaud et al. 2002). High-frequency stimulation of the motor thalamus has been shown to improve trunk spasms but not dysarthria or hypotonia (Burbaud et al. 2002); a similar technique used in the globus pallidus did not have any long-term effect (Wihl et al. 2001). Similarly, excisions in the same areas have had only partial success (Ohnishi et al. 1981; de Yebenes et al. 1988; Fujimoto et al. 1997).
Figure 1.1 - Clinical features of neuroacanthocytoses. a, Basal ganglia degeneration in ChAc. MRI scan taken in the frontal plane, showing pronounced atrophy of the heads of the caudate nuclei with subsequent widening of the lateral ventricles (black arrow), and incipient atrophy of the putamen (white arrow) (taken from Danek et al., 2004). b, Myopathy in MLS. Immunohistochemistry of quadriceps muscle stained for fast myosin, showing fibre-type grouping, fibre atrophy, centralised nuclei (arrows) and variability of fibre diameter (taken from Jung et al., 2001). c, Retinitis pigmentosa in ABL. Right retinal fundoscope of an ABL patient showing degeneration of the retinal pigment (taken from Al-Shali et al., 2003). d, ‘Eye of the tiger’ sign in HSS. T2-weighted MRI scan taken in the lateral plane, showing hypointensity (thick arrow) with a central region of hyperintensity (thin arrow) in the medial globus pallidus (taken from Hayflick et al., 2003).
Kell is a complex blood group system with more than 20 serologically recognised erythrocyte antigens. Certain Kell antigens are very immunogenic, so the Kell serotype is the next most important determinant of transfusion compatibility after ABO and Rhesus. The McLeod phenotype, characterised by weak expression of the Kell antigens and absence of the otherwise ubiquitous Kx antigen, was first recognised and named after the propositus by Allen et al. (1961). This phenotype is rare: again, no estimates of prevalence have been recorded, but Swash et al. (1983) noted that of 300,000 donors screened over three years in a British blood transfusion centre, only two with the McLeod phenotype were detected. Sixteen years after the phenotype was described, Wimer et al. (1977) discovered that the red cells in the propositus showed acanthocytosis, reduced in vivo survival and a compensated haemolytic state. Hepatomegaly and splenomegaly were subsequently observed in a New Zealand pedigree (Symmans et al. 1979). McLeod syndrome (MLS) was originally introduced to neurologists as a benign myopathy (Swash et al. 1983), but it is now appreciated that the clinical spectrum of the disorder includes severe neurological impairment, and some cases can be difficult to distinguish clinically from ChAc (Takashima et al. 1994; Stevenson and Hardie 2001; Danek 2002). In fact, some cases where neurological symptoms were initially ascribed to ChAc are now known to have MLS (e.g., Family L in (Hardie et al. 1991); case in (Faillace et al. 1982)), and it is likely that misdiagnosed cases of MLS remain in the literature.
1.2.1 **Clinical features**

The description of 22 men with the McLeod phenotype by Danek *et al.* (2001) illustrates the range of clinical features typically associated with this disorder. Liver and/or spleen enlargement was observed in 40% of the patients examined. Two-thirds of patients showed signs of cardiac disease, with a cardiac cause of death suspected in 4/5 patients. All tested patients showed areflexia or hyporeflexia, with lower limb tendons being more affected. Muscle weakness, primarily of the lower limbs, was detected in two-thirds of patients. Muscle involvement in MLS was initially recognised by Marsh *et al.* (1981), and since then has been well documented. Muscle biopsies even in subclinical cases show features of active muscle fibre necrosis and regeneration (Swash *et al.* 1983; Dotti *et al.* 2000). In addition to this myopathy, there appear to be neuropathic changes, including fibre-type grouping, in other muscle biopsies (Witt *et al.* 1992; Malandrini *et al.* 1994; Jung *et al.* 2001a) (see Figure 1.1b). Electromyographic studies usually detect the neurogenic component of the atrophy (Dotti *et al.* 2000; Danek *et al.* 2001). A muscle CT scan of three MLS patients revealed selective and symmetrical atrophy of lower limb muscles (Ishikawa *et al.* 2000). In one patient, muscle necrosis led to severe rhabdomyolysis (Jung and Brandner 2002). The muscle atrophy is reflected in laboratory findings. Elevated creatine kinase in the serum of MLS patients was first noted by Marsh *et al.* (1981) and has since been observed in nearly every patient investigated. Serum lactate dehydrogenase activity is also slightly elevated in most MLS patients (Marsh *et al.* 1981; Danek *et al.* 2001).

The age of onset of neurological features in the McLeod cohort reported by Danek *et al.* (2001) ranged from 26 - 61. Seizures were reported in 23% of patients.
Psychopathologic features, including emotional lability, anxiety and depression, inappropriate behaviour, hallucinations and paranoia, were present in at least 45% of the group. Moderate cognitive deficits were discovered in over half of those examined. The majority of MLS patients show some form of movement disorder: as in ChAc, chorea appears to be the most common manifestation, but its onset is generally later - the mean age of onset is early fifties, compared with mid-twenties for ChAc (Rampoldi *et al.* 2002). Facial tics, involuntary vocalisations and tongue movements and dysarthria have all been reported: in contrast to ChAc, though, tongue and lip biting seems to be rare. The similarities in dyskinesia between ChAc and MLS is partly explained by the similar brain pathology - MLS patients also show atrophy of the caudate nucleus by CT (Danek *et al.* 1994; Malandrini *et al.* 1994) and MRI scans show an altered signal in the caudate and putamen (Dotti *et al.* 2000). PET studies indicate that glucose metabolism is absent in the basal ganglia and low in the frontal and parietal cortex (Dotti *et al.* 2000).

### 1.2.2 Genetics

Analysis of a large New Zealand family segregating the McLeod phenotype (Symmans *et al.* 1979) proved the assumption of X-linked recessive inheritance (Wimer *et al.* 1977). Female carriers show mosaicism for expression of Kell antigens, implying that the McLeod locus is subject to X-inactivation (Marsh *et al.* 1975b; Wimer *et al.* 1977). Several boys with chronic granulomatous disease (CGD) and the McLeod phenotype had been reported (Marsh *et al.* 1975a), so the deletion of Xp21 in a patient with Duchenne muscular dystrophy (DMD), CGD, retinitis pigmentosa and MLS (Francke *et al.* 1985) implied that the gene responsible for MLS lay in this chromosomal region. Using deletion analysis of patients with different combinations...
Chapter 1: Introduction

of MLS, CGD and DMD, Bertelson et al. (1988) placed the MLS locus between the CGD and the DMD loci. Ho et al. (1992) narrowed the critical region to 150-380 kilobases within Xp21. They subsequently identified point mutations in two unrelated McLeod patients in a novel gene, \( XK \), thereby identifying it as the gene responsible for the disorder (Ho et al. 1994).

\( XK \) is arranged in three exons. Its cDNA has an open reading frame (ORF) of 1335 nucleotides, encoding a predicted protein of 444 amino acids. At time of writing, 24 different \( XK \) mutations had been reported (21 reviewed in (Rampoldi et al. 2002); (Russo et al. 2002; Jung et al. 2003; Singleton et al. 2003)). Most of these are mutations likely to lead to absence of \( XK \) protein; the mutations appear to be evenly distributed throughout the \( XK \) gene sequence (Rampoldi et al. 2002).

### 1.2.3 Possible functions of \( XK \)

The \( XK \) gene is expressed highest in foetal liver and adult skeletal muscle, slightly lower in heart, brain and pancreas, and in low levels in other tissues (Ho et al. 1994). \( XK \) protein is predicted to contain 10 transmembrane regions and 14 potential phosphorylation sites, with a large extracellular loop between the third and fourth transmembrane domains. Khamlichi et al. (1995) purified \( Kx \) as an immunocomplex with Kell, and demonstrated by amino-acid sequencing that this was identical to \( XK \). They also showed that \( XK \) is covalently attached to Kell via a disulphide bond. Interestingly, although \( XK \) is present as a complex with Kell in erythrocytes, immunohistochemistry studies have shown that in skeletal muscle, Kell localises to the cell membrane but \( XK \) is predominantly in the sarcoplasmic reticulum (Jung et al. 2001b).
Kell is a single membrane-spanning protein with a large extracellular region (Lee et al. 1991) and shows homology to mammalian endopeptidases, which process a variety of peptide hormones. Kell is primarily expressed in erythroid tissues and testis, but weak expression is seen in many other tissues (Russo et al. 2000). In vitro studies showed that Kell preferentially cleaves big endothelin-3, a peptide involved in multiple biological processes, including the regulation of vascular tone (Lee et al. 1999). Endothelins are neurotrophic factors at low concentrations and cytotoxic at high concentrations, suggesting that endothelin-related mechanisms might be implicated in neurodegeneration (Kataoka et al. 1995; Ehrenreich et al. 2000). It is therefore tempting to speculate that endothelins could play an important role in the pathogenesis of MLS. However, the simple absence of Kell and its activities on the plasma membrane cannot be the cause of McLeod syndrome, since individuals with the Kell null phenotype, lacking all Kell antigens, show no clinical or haematological abnormalities (Lee et al. 2000).

XK shows the same topographical arrangement in the membrane as Na⁺ Cl⁻ dependent neurotransmitter receptors. However, it has very little amino-acid homology with these receptors, and initial studies showed no difference between McLeod and normal red cells in uptake of various substrates (Ho et al. 1994). XK shares 19% identity and 37% similarity with CED-8 in Caenorhabditis elegans (Stanfield and Horvitz 2000). CED-8, like XK, is predicted to have 10 transmembrane regions and is located on the plasma membrane. Studies on Ced-8 loss-of-function mutations imply that the gene is involved in the execution step of programmed cell death, downstream of the caspase CED-3. The human homologue of CED-3, caspase-
8, is thought to play a crucial role in the striatal neurodegeneration in Huntington's disease (Hackam et al. 2000; Gervais et al. 2002). Stanfield et al. (2000) propose that XK may also function in apoptosis in humans and the late onset neurodegeneration associated with McLeod syndrome is a consequence of a disruption in apoptosis. However, as homology between the two proteins is not extensive, any conclusions drawn on XK’s function based on comparison with CED-8 should be treated with caution.

As McLeod myopathy bears a striking resemblance to that in female carriers of a DMD mutation (Dubowitz 1985; Witt et al. 1992), it is conceivable that dystrophin is important in the muscular manifestations of MLS. However, muscle biopsies of two MLS patients showed normal expression and distribution of dystrophin, implying that XK and dystrophin do not interact in any substantial way (Carter et al. 1990).

Clues to XK function can be provided from analysis of the amino-acid substitutions established as causal mutations in MLS. Three such missense mutations have been reported so far (Danek et al. 2001; Russo et al. 2002; Jung et al. 2003). The cysteine involved in the C294R mutation is predicted to alter the structural stability of the XK protein (Danek et al. 2001). COS-1 cells were transfected with XK cDNA encoding an R222G substitution that had been observed in a MLS patient (Russo et al. 2002). The mutant protein was expressed but did not travel to the cell surface. However, no neurological examination was reported for this patient, and so the clinical consequences of this altered XK localisation are not known. Substitution E327K has also been associated with the McLeod serological phenotype (Jung et al. 2003). Again, the mutant XK was not presented on the erythrocyte surface - but remarkably,
despite extensive investigations, no clinical or subclinical cerebral or neuromuscular signs and symptoms of MLS were found. This implies that XK physiologic function is not necessarily related to XK/Kell complex formation.

1.3 Other neuroacanthocytoses: abetalipoproteinemia

Abetalipoproteinemia (ABL, OMIM 200100) was initially described by Bassen and Kornzweig (Bassen and Kornzweig 1950), who noted an association of the signs of lipid malabsorption, acanthocytosis, pigmentary degeneration of the retina and ataxia with hypocholesterolemia. ABL is characterised by the absence of plasma lipoproteins that contain apolipoprotein B and low plasma concentrations of triglyceride and cholesterol. It may present in infancy with failure to thrive but can present in young adults as a spinocerebellar degenerative disorder with peripheral neuropathy and a pigmented retinopathy (see Figure 1.1c). About 100 cases have been reported worldwide. About a third of these cases are from consanguineous marriages: family studies suggested autosomal recessive transmission. The majority of obligate heterozygotes are reported to have normal lipid levels (Berriot-Varoqueaux et al. 2000).

1.3.1 Clinical features

The initial presentation of ABL in infancy is of lipid intolerance with steatorrhoea, vomiting and abdominal swelling. Intestine endoscopy reveals a white mucosa in the duodenum and jejunum due to storage of lipids by the enterocytes (Delpre et al. 1978). Liver biopsies consistently show hepatic steatosis (Collins et al. 1989).
Chronic malabsorption of lipids leads to lipid-soluble vitamin deficiency. If untreated, the classical neurological course of ABL includes diminution and then loss of deep tendon reflexes, then progressive alteration of proprioception, cerebellar syndrome and muscular atrophy of both neurogenic and myogenic origin (Wichman et al. 1985; Higuchi et al. 1993; Berriot-Varoqueaux et al. 2000). This atrophy leads to progressive impaired mobility. The first ophthalmological signs in ABL are alterations in night and colour vision, then a decrease in visual acuity, which can deteriorate to blindness (Berriot-Varoqueaux et al. 2000). Some patients exhibit moderate to severe anaemia due to haemolysis and shortening of erythrocyte half-life (Simon and Ways 1964). Abnormalities in coagulation have also been observed (Caballero and Buchanan 1980). At least five cases of rapidly evolving cardiac insufficiency leading to premature death have been reported (e.g., (Dische and Porro 1970)).

1.3.2 Treatment

The neurological symptoms of ABL are very similar to those in monkeys (Fitch and Dinning 1963) and rats (Machlin et al. 1977; Towfighi 1981) with vitamin E deficiency. Deficiency of vitamins A and K is believed to contribute to the visual disturbances and occasional blood coagulation problems (Caballero and Buchanan 1980; Bieri et al. 1984). Therefore, patients are typically managed by restricting dietary uptake of long-chain fatty acids, and administering supplements of essential fatty acids and vitamins A, E and K. Early diagnosis and treatment of ABL is needed to stall progression of retinal and neurological symptoms (Berriot-Varoqueaux et al. 2000).
7.3.3 Genetics

Wetterau et al. (1992) reported that a defect in the microsomal triglyceride transfer protein (MTP) might be a proximal cause of ABL. MTP is a dimeric lipid transfer protein consisting of the multifunctional protein disulfide isomerase (PDI) and a unique 97-kDa subunit. Intestinal biopsies of ABL patients revealed absence of the large subunit, although PDI was detected. Triglyceride transfer activity was undetectable. The primary involvement of MTP in ABL was confirmed later on by the cloning of the gene encoding the large subunit, MTP, and the identification of MTP mutations in patients with ABL (Sharp et al. 1993). Twenty-three MTP mutations associated with ABL have now been reported. These were reviewed by Rampoldi et al. (2002), who observed that the mutations seem to cluster in the middle and C-terminal domains of the protein. Four missense mutations have been identified - one of these, S590I, seems to be associated with a milder disease course (Wang and Hegele 2000; Al-Shali et al. 2003).

1.3.4 Function of MTP

MTP is expressed primarily in liver and intestine, although it is also observed in heart, kidney and testis. MTP has extensive amino-acid homology with vitellogenin, a transport and storage lipoprotein found in egg-laying invertebrates (Wahl 1988). A model of MTP by Mann et al., based on this similarity, predicts an N-terminal β-barrel domain, followed by a 300-aa α-helical domain, responsible for MTP interaction with PDI (Mann et al. 1999). In vitro assays have established that MTP binds and shuttles individual lipid molecules between membranes (Atzel and Wetterau 1994). The following model of apolipoprotein synthesis has been proposed (Olofsson et al. 1999). Addition of lipid to ApoB begins cotranslationally in the
endoplasmic reticulum. A small ApoB-containing lipoprotein particle is formed in the first step of lipoprotein assembly. Additional triglyceride is added in the second step assembly, probably by fusion of ApoB particle with a preformed triglyceride droplet. Cell-free binding assays, co-immunoprecipitation and yeast two-hybrid studies (reviewed in (Gordon and Jamil 2000)) suggest that initiation of lipidation of nascent apoB may occur through a direct association with MTP. This early lipidation may be required to allow the nascent polypeptide to fold properly and therefore avoid targeting by Hsp70 for proteasome degradation. Experiments with transgenic animals support a primary role for MTP in ApoB lipid loading. Homozygous MTP knockout mice die in midgestation, suggesting a rather different role of MTP in rodent embryological development (Raabe et al. 1998), but conditional knockout of hepatic MTP leads to complete loss of plasma ApoB100 (Chang et al. 1999; Raabe et al. 1999).
1.4 Other neuroacanthocytoses: hypobetalipoproteinemia

Familial hypobetalipoproteinemia (HBL, OMIM 107730) is a codominant disorder. Heterozygotes have plasma levels of cholesterol and ApoB below the fifth percentile and are typically asymptomatic with no acanthocytes (Schonfeld 2003), although recent reports reveal hepatic steatosis in a large proportion of heterozygotes (Tarugi and Lonardo 1997; Schonfeld et al. 2003). The frequency of HBL heterozygotes is estimated to be 1/500 - 1/1000 in Western populations (Linton et al. 1993). Homozygotes have extremely low levels of apoB; clinically they are very variable and severe cases are indistinguishable from ABL, with neuro-ophthalmological complications and acanthocytes (Herbert et al. 1983). Vitamin E treatment is also standard for these cases, but, as in ABL, doesn’t reverse the acanthocytosis (Hegele and Miskie 2002).

1.4.1 Genetics

Young et al. (1987; 1988) demonstrated that mutations in the APOB gene on chromosome 2 cause some cases of HBL. Over 40 mutations have now been described (reviewed in (Schonfeld 2003)). Most of the reported mutations are nonsense or frameshift mutations that prevent translation of full-length apoB-100. Seven families with no linkage to the APOB locus show evidence of linkage to chromosome 3p21 (Yuan et al. 2000; Neuman et al. 2002), but other families show linkage to neither locus (Schonfeld 2003) and their genetic lesion is yet to be determined.
1.5 Other syndromes associated with acanthocytosis

1.5.1 Hallervorden-Spatz syndrome

The term ‘Hallervorden-Spatz syndrome’ (HSS) applies to a spectrum of disorders that share the common features of neurodegeneration and iron accumulation in the brain. A characteristic MRI finding in T2-weighted images is decreased signal intensity from the globus pallidus with a central area of increased signal (the ‘eye of the tiger’ sign, see Figure 1.1d). Classic HSS is an autosomal recessive disorder that presents in childhood with dystonia, dysarthria and rigidity and leads to mental deterioration and an early death. Some patients show retinal degeneration (Hayflick 2003). Atypical HSS has a later onset and a more slowly progressive course.

Mutations in the \textit{PANK2} gene, encoding pantothenate kinase 2, are responsible for all classic HSS and a third of atypical HSS (Zhou \textit{et al.} 2001; Hayflick \textit{et al.} 2003): consequently, this disorder is now known as pantothenate kinase-associated neurodegeneration (PKAN, OMIM 234200). Acanthocytosis was first associated with PKAN by Roth \textit{et al.} (1971), and was reported in eight percent of a large cohort of PKAN patients with classic symptoms (Hayflick \textit{et al.} 2003). Since acanthocytosis is not routinely sought in PKAN, its true prevalence in this disorder may be higher.

PKAN mutations have also been found in patients with the eye of the tiger sign and acanthocytosis, retinitis pigmentosa and hypoprebetalipoproteinemia (HARP syndrome) (Ching \textit{et al.} 2002; Houlden \textit{et al.} 2003). Pantothenate kinase is a key regulatory enzyme in the biosynthesis of coenzyme A, which plays a central role in intermediary and fatty acid metabolism (Abiko 1967). Interestingly, the \textit{Drosophila} pantothenate kinase hypomorphic mutant shows impaired neurologic co-ordination (Afshar \textit{et al.} 2001).
1.5.2 Huntington’s disease-like 2

Walker et al. (2002) described a pedigree with autosomal dominant inheritance of a progressive neurodegenerative disorder comprising chorea, dementia, parkinsonism and acanthocytosis. Based on its similarities to the ChAc phenotype, they called the disorder ‘autosomal dominant chorea-acanthocytosis’. The affected individuals in this family were subsequently found to harbour the trinucleotide repeat expansion within the gene JPH3 that has been associated with Huntington’s disease-like 2 (HDL2) (Holmes et al. 2001; Walker et al. 2003). In light of these findings, the blood of six additional HDL2 patients was screened: and one showed acanthocytes (Walker et al. 2003). JPH3 codes for junctophilin-3, a component of the junctional complex that anchors plasma membrane to endoplasmic reticulum, and that may be involved in the functional coupling between cell-surface voltage sensors and intracellular calcium channels (Nishi et al. 2000; Takeshima et al. 2000). Junctophilin-3 knockout mice show impaired motor co-ordination (Nishi et al. 2002).
1.6 The basis of acanthocytosis

1.6.1 Erythrocyte structure

The red blood cell membrane is composed of a lipid bilayer to which is anchored a filamentous network of proteins that underlie the cytoplasmic side of the membrane (reviewed in Marchesi et al. 1976; Mohandas and Gascard 1999; Gascard and Mohandas 2000); see Figure 1.2a). About 52% of the membrane mass is protein, 40% is lipid and 8% is carbohydrate. The phospholipid bilayer provides physical continuity to the membrane and serves as the matrix in which transmembrane proteins reside. The membrane skeleton consists of several protein components, the most abundant of which is spectrin. Spectrin is a flexible, rod-like molecule. It is a heterodimer, with the subunits intertwined side-to-side. The heterodimers associate head-to-head to form tetramers; the tail ends associate with short oligomers of actin, in a ternary complex with protein 4.1R that is further stabilised by interaction with adducin. This membrane skeleton is principally linked to the lipid bilayer by ankyrin, which associates with the transmembrane band 3 protein and spectrin; and by the binding of glycoporphin C in the bilayer to 4.1R. This composite protein structure plays a critical role in regulating cell shape, mechanical properties and structural stability. The ability of the red cell to undergo extensive deformation is essential, not only for its function (its diameter, 8 μm, far exceeds that of the capillaries, 2-3 μm), but also for its survival (Mohandas and Chasis 1993).
Figure 1.2 - Normal and abnormal erythrocyte structure. **a**, Organisation of the erythrocyte plasma membrane. Ankyrin and protein 4.1R bind glycophorin C (GPC) and band 3, respectively, to the spectrin/actin cytoskeleton. Additional proteins, including protein 4.2, protein p55, adducin, tropomyosin and tropomodulin, participate in regulation of these interactions (adapted from Gascard and Mohandas, 2000).

**b**, Acanthocytosis. Scanning electron microscopy of peripheral blood from a ChAc patient reveals many stage I (aI) and stage II (aII) acanthocytes, together with a few morphologically normal discocytes (d) (taken from Danek et al., 2004).
1.6.2 Normal and aberrant erythrocyte morphology

Normal erythrocyte shape is that of a biconcave disk, the discocyte. The bilayer couple hypothesis of red cell shape, proposed by Sheetz and Singer (1974), states that the asymmetric expansion of the outer leaflet leads to spiculated forms such as acanthocytes and echinocytes, while expansion of the inner leaflet leads to formation of stomatocytes (rounded cells with internal invaginations). Acanthocytes are defined as dense, slightly contracted red cells with a number of irregularly spaced, thorny surface protrusions, some with terminal bulbs (Brecher and Bessis 1972); see Figure 1.2b). Echinocytes have more abundant and evenly distributed surface projections that have a much broader base in relation to length. The distinction between acanthocytes and echinocytes is important, as healthy individuals may have up to 3% echinocytosis in a peripheral blood smear (Dacie 1985). In contrast, acanthocytes are not normally found in the blood. Unlike echinocytes, the formation of acanthocytes cannot be readily induced. Studies on ABL patients show that the in vivo deformation of the membrane is irreversible and produced after erythrocytes pass into the peripheral circulation (Bassen and Kornzweig 1950).

1.6.3 Detection of acanthocytes

Reported percentages of acanthocytes in affected individuals’ blood vary widely. Patients with ABL usually have over 50% acanthocytosis (Berriot-Varoqueaux et al. 2000). In contrast, patients with ChAc or McLeod syndrome are usually in the 5-50% range (Hardie et al. 1991). There appears to be no correlation between degree of acanthocytosis and severity of symptoms or length of disease course. In addition, some patients with molecularly confirmed ChAc have exhibited no acanthocytes during the disease course (Malandrini et al. 1993; Johnson et al. 1998), or only many
years after the onset of neurological symptoms (Sorrentino et al. 1999). It is not clear how much variability in acanthocytosis can be accounted for by differing laboratory procedures. For highest sensitivity of acanthocyte detection, wet film preparations of blood diluted 1:1 with heparinised saline have been recommended (Feinberg et al. 1991). However, routine haematological investigations typically utilise dried blood smears of whole blood, which were not sensitive enough to determine abnormalities in 2/3 patients with confirmed ChAc (Storch et al. 2004).

1.6.4 Factors in acanthocyte formation

ChAc blood cells that are initially morphologically normal show a dramatic increase in sensitivity to stresses that induce echinocytic change (Feinberg et al. 1991; Terada et al. 1999). ChAc red blood cells exhibit an increased density (Clark et al. 1989; Terada et al. 1999); and reduced deformability (Villegas et al. 1987; Clark et al. 1989), possibly due to decreased membrane fluidity (Oshima et al. 1985; Villegas et al. 1987). Since the major components of cell membrane are lipid, protein and carbohydrate, abnormalities in any one of these must be considered. The carbohydrate moiety is not known to be involved in morphology of red cells (Galey et al. 1978). Instead, studies on the formation of acanthocytes have focussed on the lipid and protein contributions.

1.6.5 Role of lipids

The outer and inner leaflets of the normal erythrocyte membrane differ in lipid composition (Rothman and Lenard 1977), and the lipid fluidity of the outer leaflet exceeds that of the inner (Cogan and Schachter 1981). In ABL, it is generally believed that an altered lipid composition is responsible for the acanthocytosis. Global amounts
are little changed, but the ratio of sphingomyelin to lecithin is always increased. The phospholipid fatty acid composition is deficient in linolenic and arachidonic acids, and increased in saturated fatty acids (Berriot-Varoqueaux et al. 2000). This alteration produces a decrease in fluidity (Cooper 1977), which Flamm and Schachter (1982) demonstrated was restricted to the outer leaflet of the erythrocyte membrane. The imbalance in lipid composition is probably related to the absence of betalipoprotein in the plasma. Normal erythrocytes acquire an acanthocytic form when transfused into ABL patients (Frezal et al. 1961). However, in vitro incubation of normal red cells with ABL plasma doesn’t induce this change (Schwartz et al. 1961; Cooper and Gulbrandsen 1971), which suggests the deformation is not due to a simple passive exchange of lipids between erythrocyte and plasma.

In contrast, studies on the lipid contribution to acanthocytosis in ChAc have been more equivocal. Sakai et al. (1991) noted a decrease in linolenic acid in ChAc erythrocytes, and abnormal composition of the fatty acids covalently bound to membrane proteins. Clark et al. (1989) enriched the acanthocyte population in ChAc blood by density-gradient centrifugation and found an increase in the ratio of sphingomyelin to phospholipids. However, other studies have reported no significant differences in red blood cell membrane lipid or fatty acid composition (Villegas et al. 1987; Hardie et al. 1991; Olivieri et al. 1997). Oshima et al. (1985) also found no definite abnormalities in the phosphatide content, or cholesterol/phospholipid ratio, so suggested that distortion of the membrane in ChAc acanthocytes is caused by factors other than the lipid components.
1.6.6 Role of proteins

Proteins are known to play a role in membrane fluidity (Cooper et al. 1980). It is therefore possible that a dysfunction of lipid-protein or protein-protein interaction could lead to acanthocytosis. A freeze-fracture study of ChAc erythrocyte membranes showed a marked increase of intramembranous particle-free areas (Ueno et al. 1982). This was not replicated by Clarke et al. (1989), but recent studies have demonstrated focal changes in the compactness of membrane protein networks in ChAc erythrocytes (Terada et al. 1999). Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) of freshly prepared erythrocyte membranes shows no gross abnormalities in ChAc protein composition or mobilities (Feinberg et al. 1991; Bosman et al. 1994). However, SDS-PAGE of older preparations indicate that ankyrin and bands 3 and 4.2 are self-digested faster in ChAc patients than in healthy controls, possibly due to a conformational defect in the junction between spectrin and band 3 (Asano et al. 1985). Western blotting also shows increased fragmentation of band 3 in some patients (Bosman et al. 1994).

There is additional evidence for the involvement of band 3 in acanthocytic change. Band 3 is a ubiquitous protein. In red blood cells, it mediates the exchange of anions across the membrane (Cabantchik and Rothstein 1974; Lepke et al. 1976), maintains acid-base balance and is the binding site for several glycolytic enzymes and haemoglobin (Strapazon and Steck 1977; Salhany and Shaklai 1979; Kliman and Steck 1980). Band 3 undergoes an as-yet uncharacterised initial change during cellular ageing. Following this, band 3 degrades and becomes ‘senescent cell antigen’. A physiologic IgG autoantibody binds to this and initiates cellular removal (Kay 1981). Bruce et al. (1993) discovered that two siblings showing acanthocytosis but no
other clinical features were homozygous for a P896L mutation in band 3 protein, which increased its apparent electrophoretic mobility and anion transport and decreased the number of ankyrin-binding sites (Kay et al. 1994). ChAc erythrocytes show immunologic characteristics of aged band 3, and antibodies to brain anion transporter have been isolated from the sera of ChAc patients (Kay et al. 1990a; Kay et al. 1990b). Kay et al. propose that an alteration in conformation in band 3 in ChAc patients leads both to acanthocytosis in erythrocytes, and premature destruction of neurons in the basal ganglia. How directly or indirectly the band 3 alteration is linked to the primary genetic defect in ChAc is not yet clear.

The structural defect of McLeod syndrome red cells may also lie in the protein composition as there appears to be no abnormality or deficiency in McLeod erythrocyte phospholipids (Galey et al. 1978; Kuypers et al. 1985), and they show normal transport of phospholipids (Redman et al. 1988). Dallas et al. (1990) demonstrated that MLS red cells showed decreased deformability, partly due to a reduction in surface area and partly to intrinsic membrane stiffness. The obvious primary protein candidates for structural changes in McLeod erythrocytes are XK and Kell, since they are both transmembrane proteins whose presence is abolished or greatly reduced in this disorder. As previously discussed (section 1.2.3), the absence of Kell cannot be a primary cause for acanthocytosis, since Ko individuals have morphologically normal erythrocytes. Suspicion has therefore rested on the XK protein. However, apart from Kell, XK was not shown to associate with any transmembrane or membrane-associated proteins (Russo et al. 1998), and certain individuals in whom membrane XK was undetectable did not show acanthocytosis.
(Jung et al. 2003). The exact basis for acanthocytic change in MLS, as in ChAc, is therefore still to be elucidated.

### 1.7 Genetics of ChAc

#### 1.7.1 Mode of inheritance

The mode of transmission of ChAc was originally unclear. A few kindreds with apparent autosomal dominant inheritance have been reported (Levine et al. 1968; Kito et al. 1980; Marson et al. 2003; Saiki et al. 2003), but no studies have shown unequivocal evidence of dominant transmission. Ironically, it is possible that the original family reported by Levine et al. (1968) was a manifestation of McLeod syndrome with skewed female X-inactivation (R Hardie, personal communication). In any case, the vast majority of familial cases that have been reported show ChAc in only one generation, and there are a relatively high number of cases from consanguineous parents (e.g., (Bird et al. 1978; Sakai et al. 1981; Yamamoto et al. 1982; Sotaniemi 1983; Spitz et al. 1985; Serra et al. 1986; Alonso et al. 1989; Tanaka et al. 1998)), implying that the ‘classic’ ChAc neurological phenotype is an autosomal recessive trait. In contrast, semidominant inheritance of acanthocytosis has been observed in some pedigrees (Sotaniemi 1983; Dubinsky et al. 1989; Ueno et al. 2001).

#### 1.7.2 CHAC positional cloning

Rubio et al. (1997) performed a genome-wide linkage screen using 11 families of diverse geographical origin segregating for ChAc. The disorder was linked in all
families to a 6-centimorgan interval on chromosome 9q21-q22, flanked by the recombinant markers $GATA89A11$ and $D9S1843$ (see Figure 1.3a). This result was confirmed by homozygosity-by-descent analysis in offspring from consanguineous marriages, and indicated that ChAc is indeed an autosomal recessive disorder. A partial yeast artificial chromosome contig was constructed and 17 transcripts were mapped in the ChAc critical region. Two of these, $GNAQ$ and $GNA14$, encode members of the guanine nucleotide-binding protein $\alpha$-subunit family (Simon et al. 1991). $GNAQ$ and $GNA14$ were screened for mutations in ChAc patients but no significant changes were found (Rubio et al. 1999).

We subsequently identified a novel gene, denoted $CHAC$, in this critical region. Direct sequencing of probands from the 11 families used for linkage analysis revealed 16 different $CHAC$ mutations likely to disrupt or abolish gene function (Rampoldi et al. 2001). Another group who independently identified the gene reported one additional mutation, a deletion of exons 60 and 61, in three families from the same region in Japan (Ueno et al. 2001). $CHAC$ has recently been renamed $VPS13A$, to acknowledge its similarity with $VPS13/SOI1$ in *Saccharomyces cerevisiae* (Velayos-Baeza et al. 2004).

1.7.3 Genomic structure of $VPS13A$

$VPS13A$ is organised in 73 exons, spanning about 240 kilobases of the genome, as shown in Figure 1.3. It has multiple splice forms (Velayos-Baeza et al. 2004), but there are two predominant transcripts as determined by expressed sequence tag (EST) analysis and reverse-transcription polymerase-chain reaction (RT-PCR). Transcript 1A contains exons 1 to 68 and 70 to 73, therefore skipping exon 69, which has an in-
Figure 1.3 - Physical map of the CHAC region on chromosome 9q21-q22 and genomic structure of VPS13A. a, Schematic of the CHAC critical region defined by the recombinant markers (bold) GATA89A11 and D9S1843. The localisation of polymorphic markers (bold) and known genes is shown. Positions of markers and genes are derived from the UCSC Genome Browser July 2003 assembly (http://genome.ucsc.edu/cgi-bin/hgGateway). b, Intron/exon organisation of VPS13A. Exons are indicated as thick bars (white blocks represent 5' and 3' UTRs). Introns are depicted as thin lines and are not scaled. The start ATG is located at position 260 in the transcript sequence. The two major splicing variants 1A (exons 1-68 + 70-73) and 1B (exons 1-69) are shown. The locations of polyadenylation signals in exon 69 and exon 73 are indicated by triangles.
frame stop codon. It has a full-length sequence of 11262 bp containing an ORF of 9525 nt and a 3' untranslated region (UTR) sequence of 1477 bp. Analysis of ESTs and cDNA clones suggests the preferential use of one consensus and one alternative polyadenylation signal at positions 10863 and 9715 (from the initiating ATG), respectively. Transcript 1B comprises exons 1-69. It is 10050 bp long, containing an ORF of 9288 nt. This encodes a protein of 3095 aa which lacks the 111 aa encoded by exons 70 to 73 and has 32 new amino acids at the C-terminus. The 500 bp 3' UTR contains one consensus polyadenylation signal at position 9776. The protein encoded by these transcripts is called chorein.

1.8 Chorein

Little is known of chorein's structure or function. Chorein is annotated in the SWISSPROT database as having 10 tetratricopeptide repeats (TPRs). The TPR is a loosely conserved 34-aa motif widespread in evolution, with eight residues conserved in terms of size, hydrophobicity and spacing (Lamb et al. 1995; Blatch and Lassle 1999). TPR-containing proteins are involved in a number of processes, including cell-cycle control, stress response, mitochondrial and peroxisomal protein transport and neurogenesis (Goebl and Yanagida 1991): there appears to be no common biochemical function linking these proteins. TPRs are predicted to form two amphipathic helices that can mediate TPR-TPR interactions (Hirano et al. 1990; Sikorski et al. 1990), although many interactions between TPR and non-TPR-containing proteins have been reported (Blatch and Lassle 1999). Computer searches of chorein's sequence do not identify any other known structural motifs or domains,
and the number of putative transmembrane domains in chorein diverges widely between different prediction programs. Initial clues to chorein's function must therefore come from analysis of its protein homologues.

1.8.1 Chorein orthologues

Chorein seems to belong to a novel gene family, as shown in Figure 1.4. It is very highly conserved (83% identities, 90% positives) in mouse. It has a lower degree of sequence similarity (20-29% identities, 39-48% positives) with five proteins of comparable size: the putative protein products T08G11.1 and CG2093, of *Caenorhabditis elegans* and *Drosophila melanogaster* respectively, the protein Vps13p of *Saccharomyces cerevisiae* and two orthologues in *Schizosaccharomyces pombe*. The sequence similarity significantly increases (up to 49% identities for the *Drosophila* and *C. elegans* proteins) in the amino (N-) and carboxy (C-) termini, where matches can also be found with a VPS13-like protein of *Arabidopsis* and the protein TipC of *Dictyostelium*. As these proteins are of an unusually large size and show strong conservation of the terminal domains, it is likely that they have all descended from a common ancestor.
Figure 1.4 - Putative chorein orthologues. Schematic representation of the homology of hypothetical (italics) and known proteins with chorein. Chorein regions 1-4 (indicated at top) were separately analysed using the NCBI pairwise blastp program (http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html). Numbers within each region correspond to degree of sequence similarity with chorein. Only those regions with significant similarity are shown. Dashed lines connect regions originating from the same protein; different products believed to comprise part of the same protein, as based on physical map positions, are shown on the same line. 25/44, 25% identical/44% positive amino acids; Hs, Homo sapiens; Mm, Mus musculus; Ce, Caenorhabditis elegans; Dm, Drosophila melanogaster; Sc, Saccharomyces cerevisiae; Dd, Dictyostelium discoideum; Sp, Schizosaccharomyces pombe; At, Arabidopsis thaliana; orth, orthologue.
Of the proteins for which experimental evidence is available, the closest orthologue of chorein is Vps13p in *S. cerevisiae*. In yeast, *vps* genes encode proteins required for proper trafficking from the Golgi to the vacuole (Banta *et al.* 1988; Raymond *et al.* 1992). Vps13p (also reported as Soi1p) is a highly charged, 3144-aa protein (Redding *et al.* 1996). It appears to promote the cycling of the *trans*-Golgi network (TGN) transmembrane proteins Kex2p, Vps10p and Ste13p, between the TGN and the prevacuolar compartment (PVC) (Redding *et al.* 1996; Brickner and Fuller 1997). The PVC corresponds to the multivesicular body/late endosome in mammalian cells.

Vps13p antagonises a retention signal at the TGN and, at the PVC, promotes the entry of proteins containing a retrieval signal in retrograde transport vesicles. It appears to exist in a high molecular weight heterooligomeric or homooligomeric complex peripherally associated with membranes, and is present at very low levels in cells. Brickner *et al.* note that Vps13p’s highly conserved domains at the N- and C-termini, with a monotonous amino acid composition rich in aliphatic and charged residues in between, are reminiscent of intermediate filament proteins (Fuchs and Weber 1994), and suggest that it may form a cytoskeletal element that physically links the TGN and PVC. Vps13p has subsequently been implicated in the trafficking of insulin via its interaction with Kex2p (Zhang *et al.* 2001).

In nonhumans, the only other significantly similar known protein is TipC. The *tipC* gene encodes a 3848-aa protein that is implicated in *Dictyostelium* development (Stege *et al.* 1999). When exogenous nutrients are depleted, *Dictyostelium* amoebae stop growing and aggregate to produce a multicellular mound. A critical process in the morphogenesis of *Dictyostelium discoideum* is the formation of the tip at the apex of the mound that extends after an extracellular matrix surrounds each aggregate.
tipC mutants aggregate normally to the mound stage, but instead of forming a single tip from each mound as in wild type, the mutants extend multiple tips (Stege et al. 1999). The authors suggest that the tipC mutants have some defect in the cell sorting that enables morphogenesis beyond the aggregate stage.

1.8.2 The human VPS13 family

We have recently characterised the structure of three new human genes encoding proteins similar to chorein, located on chromosomes 1p36, 8q22 and 15q21 (Velayos-Baeza et al. 2004) (see Figure 1.5). Like chorein, their closest yeast homologue is also Vps13p: we have therefore named the genes VPS13B (on chromosome 8q22), VPS13C (15q21) and VPS13D (1p36). All are large genes with a high number of exons spanning 208 - 864 kilobases of genomic sequence. VPS13C is most similar to VPS13A and may have arisen by recent gene duplication: intragenic duplication of multiple exons also seems to be a common feature of these genes. Each gene is expressed ubiquitously and has at least two major alternatively spliced forms: interestingly, RT-PCR analysis revealed a brain-specific splice-form expression pattern for all genes except VPS13A. Computer analysis predicts that VPS13D contains an ubiquitin-associated (UBA) domain, present in some proteins involved in the ubiquitination pathway; and that a VPS13C splicing isoform contains a lysine-rich motif. Neither domain is predicted in chorein.
VPS13B has recently been reported as COHI: eight different mutations leading to a premature termination codon (PTC) and one missense mutation in COHI were found in 31 patients with Cohen syndrome (Kolehmainen et al. 2003). This disorder presents in childhood: its clinical characteristics include a nonprogressive psychomotor retardation, microcephaly, hypotonia and joint laxity, retinochoroidal dystrophy, myopia, and neutropenia (Kivitie-Kallio and Norio 2001). The limited similarity between the phenotype of Cohen syndrome and ChAc patients is perhaps reflected in the fact that of the four human VPS13 proteins, chorein and VPS13B share the least homology, mainly restricted to the N- and C-termini.
Figure 1.5 - The human VPS13 family. Schematic representation of the homology of human VPS13 proteins with chorein. Chorein regions 1-4 (indicated at top) were separately analysed using the NCBI pairwise blastp program (http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html). Splice variant 1A was used for each protein. Numbers within each region correspond to degree of sequence similarity with chorein. Dashed lines connect regions originating from the same protein. VPS13C and VPS13D show overlapping alignments against region 2 of chorein: this probably reflects intramolecular duplications that have occurred in the evolution of these two proteins. Key: 62/77, 62% identical/77% positive amino acids.
1.9 Research strategy

This thesis describes the mutations identified in 11 ChAc families with linkage to chromosome 9q21-22, thus demonstrating that mutations in \textit{VPS13A} are responsible for chorea-acanthocytosis. \textit{VPS13A} encodes a large, evolutionarily conserved protein called chorein. Other than the assumption that chorein plays a similar role to its structural homologue in \textit{Saccharomyces cerevisiae}, little can be said of its potential function or how its absence leads to the ChAc phenotype. As the first step towards addressing these issues and to obtain an idea of the mutational spectrum of this disorder, an additional 72 unrelated probands with ChAc were screened for \textit{VPS13A} mutations (Chapter 3).

\textit{VPS13A} expression was investigated at the mRNA level by Northern blot and RT-PCR analysis. Polyclonal antisera raised against an N-terminal region of the protein were subsequently generated. The antisera were used for immunoblot detection in a range of tissues, in order to investigate the normal distribution of chorein expression, and to determine the effect of various \textit{VPS13A} mutations on chorein expression in ChAc patients (Chapter 4).

As the molecular pathways involved in chorein function and pathogenesis of ChAc are completely unknown, a yeast-two hybrid screen was used to search for potential chorein-interacting proteins. In order to verify the interaction and determine its biological significance, co-immunoprecipitation and immunolocalisation studies were performed on a putative chorein-binding partner (Chapter 5).
Chapter 2: Materials and Methods

This chapter describes the techniques used in this study. Reagents were purchased from Sigma, unless otherwise stated. All enzymatic reactions were set up on ice, and all other procedures were performed at room temperature, unless otherwise stated. Reagent compositions not detailed in the text are listed in Tables 2.1 - 2.3.

2.1-2.10 NUCLEIC ACID TECHNIQUES

2.1 Preparation of nucleic acids

2.1.1 Extraction of genomic DNA from blood
Genomic DNA (gDNA) was extracted from 5-10 ml peripheral blood using the Nucleon® Genomic DNA Extraction Kit BACC2 (Tepnel Life Sciences, plc), according to manufacturer’s instructions. Briefly, erythrocytes were lysed by incubation in a high-sucrose solution. The lymphocytes were pelleted and then lysed in buffer containing 1% sodium dodecyl sulphate (SDS). Samples were deproteinised by addition of sodium perchlorate, before extraction of proteins by treatment with chloroform and Nucleon resin. DNA was precipitated by addition of 2 volumes of cold absolute ethanol. The DNA ‘jellyfish’ precipitate was removed with an inoculating loop, briefly washed in 70% ethanol and allowed to air dry, before resuspension in 300-600 µl 0.1 x TE, pH 8.0. If no jellyfish was visible, the mixture was kept at -20°C for ≥ 2 h, then subjected to centrifugation at 2200g for 30 min. The pellet was washed by addition of 70% ethanol and centrifugation at 2200g for 5 min.
Chapter 2: Materials and Methods

The washed pellet was allowed to air dry and was then resuspended in 300 μl 0.1 x TE, pH 8.0.

2.1.2 Extraction of gDNA from lymphoblastoid/fibroblast cell lines

Lymphoblastoid cells were grown to maximal density in a 75-cm² tissue culture flask. The cells were then pelleted and washed, as described in section 2.11.4. Fully confluent fibroblasts were trypsinised as described in section 2.11.3 before pelleting. Washed cell pellets were resuspended by vortexing in 5 ml transport buffer containing 0.2 mg/ml proteinase K, and incubated with shaking at 55°C overnight.

Deproteinisation and all subsequent steps were performed (with scaled-up volumes) according to the Nucleon BACC2 kit instructions. Due to their high protein content, cell digests were often extracted twice with chloroform before precipitation of DNA.

2.1.3 Preparation of recombinant DNA from bacteria - minipreps

Escherichia coli transformed with high-copy plasmids were cultured overnight in 4 ml Luria Broth (LB). 1 ml was removed for preparation of glycerol stocks (section 2.10.5). The remaining 3 ml were processed using the QIAprep Spin Miniprep kit (QIAGEN), according to manufacturer’s instructions. Plasmid DNA was eluted in 50 μl 10 mM Tris.Cl buffer, pH 8.5.

2.1.4 Preparation of recombinant DNA from bacteria - maxipreps

E. coli transformed with yeast two-hybrid cDNA library constructs (section 2.21.4) were cultured overnight in 50 ml Terrific Broth. The overnight culture was processed using the QIAGEN Plasmid Maxi Kit and a QIAGEN-tip 500 column, according to manufacturer’s instructions. Plasmid DNA was resuspended in 200 μl TE, pH 8.0.
2.1.5 Extraction of RNA

Total RNA was extracted from cultured cells using TRIzol® Reagent (Invitrogen), following the manufacturer’s instructions. RNA from 1-4 x 10^6 cells was resuspended in 60 µl RNase-free water.

2.2 Polymerase chain reaction (PCR)

All PCRs were carried out on a PTC-225 thermal cycler (MJ Research). Primer sequences, and the corresponding MgCl₂ concentrations and annealing temperatures (ATs) used, are listed in Appendix 1.

2.2.1 Standard PCR

Standard reactions contained 1x PCR buffer, 1-4 mM MgCl₂ (both from Applied Biosystems), 0.2 pmol/µl each forward and reverse primer, 0.2 mM dNTPs and 0.02 U/µl AmpliTaq Gold polymerase (Applied Biosystems). The PCR template was 32-40 ng gDNA, unless otherwise stated. Reactions were typically conducted in a total volume of 50 or 100 µl. Amplification conditions were typically as follows: initial denaturation and polymerase activation step of 95°C for 18 min; 30 cycles of denaturation at 95°C for 30 s, AT for 30 s, elongation at 72°C for 30 s; final elongation at 72°C for 10 min. For amplification of 500-bp to 1-kb fragments, the elongation cycle step was increased to 1 min.
2.2.2 **PCRs for denaturing high-performance liquid chromatography analysis**

When amplifying gDNA for mutation screening by denaturing high-performance liquid chromatography (DHPLC) (section 2.8.2), care had to be taken to reduce PCR error. Reactions were set up as above, except that a proofreading polymerase mix was used (9:1 volume mix of AmpliTaq Gold polymerase: *PfuTurbo* DNA Polymerase (Stratagene)), at 0.2 μl per 50 μl reaction. To reduce unspecific primer annealing, a ‘touchdown’ program was used. Thermocycling was carried out as follows: 95°C for 15 min; a touchdown of 14 cycles (95°C for 30 s; AT + 7.5°C with -0.5°C per cycle for 30 s; 72°C for 30 s); 25 standard PCR cycles; 72°C for 7 min. Amplification of GC-rich *VPS13A* exon 1 was carried out as above, except that deaza dGTP-dNTP mix was substituted for standard dNTPs, and thermocycling was performed on a GeneAmp® PCR System 9700 (Applied Biosystems).

2.2.3 **Long template PCR**

Amplification of 1.2-2 kb fragments from 3 μl lymphoblastoid first-strand cDNA (section 2.2.5) was performed using *rTth* DNA polymerase XL (Applied Biosystems), according to manufacturer’s instructions. Thermocycling was carried out as follows: 94°C for 2 min; 34 cycles of 94°C for 15 s, 60°C for 30 s, 72°C for 2 min; 72°C for 10 min. Amplification of 3-6 kb fragments from 40 ng gDNA was performed using BIO-X-ACT™ Long DNA Polymerase (Bioline Ltd), according to manufacturer’s instructions. Thermocycling was carried out as follows: 95°C for 5 min; AT for 1 min; 30 cycles of 68°C for 4-7 min, 95°C for 30 s, AT for 30 s; 68°C for 10 min.
2.2.4 **Colony PCR**

This technique was used to screen large numbers of bacterial colonies for the construct of interest. Colonies were suspended in 50 µl dH₂O. 5 µl of this suspension was used as template in a standard PCR reaction, scaled down to 25 µl total volume. The remaining colony suspension from a PCR-positive clone was used to inoculate an overnight culture for subsequent miniprep (section 2.1.3).

2.2.5 **Reverse-transcription PCR (RT-PCR)**

Total RNA (5 µg), diluted in water to 29.5 µl total volume, was denatured at 70°C for 10 min then cooled on ice. First strand cDNA was synthesised by reverse transcription in a 50 µl reaction containing 100 ng random hexanucleotides (Amersham Biosciences), 40 U RNAsine Ribonuclease Inhibitor (Promega), 1 mM dNTPs, 1 x first strand buffer, 10 mM dithiothreitol (DTT) and 200 U Superscript RNAse H Reverse Transcriptase (all from Invitrogen). Reactions were incubated at 37°C for 45 min, and were then stored at -20°C until needed. Standard 25-µl PCRs (section 2.2.1) or long-template PCRs (section 2.2.3) were performed using 2.5 µl RT reaction as template. To analyse tissue distribution of *VPS13A* mRNA, Human Multiple Tissue cDNA (MTC™) Panel 1 (CLONTECH Laboratories, Inc) was used as template - 2 µl were used in a 25-µl PCR.

2.3 **Agarose gel electrophoresis**

DNA molecules were resolved according to size using standard agarose gel electrophoresis. 2% agarose gels were typically used, but for separation of molecules
greater than 1 kb, 1% gels were used. For restriction analysis of <500-bp fragments, 3-4% gels were used. Agarose was melted in 1 x Tris-acetate EDTA (TAE) buffer (Bio-Rad Laboratories). After cooling to below 50°C, ethidium bromide was added to 0.4 mg/ml. Samples were mixed with agarose loading buffer before being loaded on the gel and subjected to electrophoresis at 110-160 mA. 1Kb PLUS DNA ladder (Invitrogen) was used as size standard. Analytical gels were visualised and photographed using an AlphaImager™ 1220 (Flowgen). Preparative gels were visualised using a UV transilluminator; the DNA fragment of interest was excised and purified as in section 2.5.

2.4 Quantification of DNA and RNA

2.4.1 Spectrophotometric quantification

DNA and RNA were quantified by measuring the optical density at 260 nm (OD$_{260}$) of appropriate dilutions on a DU 640 Spectrophotometer (Beckman Coulter Bioresearch), or of undiluted samples on the NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies, Inc). Concentrations were calculated according to Sambrook et al. (1989):

DNA concentration (ng/µl) = OD$_{260}$ x dilution x 50

RNA concentration (ng/µl) = OD$_{260}$ x dilution x 40
2.4.2 Quantification on agarose gel

Estimates of DNA yield from PCRs and post-digest purifications were made after electrophoresis alongside known quantities of HindIII-digested λ DNA (Invitrogen) and comparison of band intensities.

2.5 PCR product/restriction digest purification

Column purification of all DNA >100 bp from a reaction was achieved using the QIAquick PCR Purification Kit (QIAGEN), according to manufacturer's instructions. If multiple samples were to be purified, the 96-well MultiScreen™ MANU 030 PCR purification plate (Millipore) was used. Samples were added to the plate, which was attached to a vacuum manifold until wells became dry. Samples were eluted by addition of dH₂O to wells, and incubation with shaking for 10 min. DNA from a 50-µl PCR would typically be eluted in 30 µl dH₂O. Purification of specific DNA fragments from agarose gel was performed with the GENECLEAN® SPIN kit (Qbiogene), according to manufacturer’s instructions.
2.6 Nucleic acid hybridisation protocols

2.6.1 Southern blot detection

Transfer. Genomic DNA (8 μg) digested with restriction endonuclease was subjected to electrophoresis on a 1% agarose gel. The gel was photographed alongside a ruler to enable accurate sizing of DNA fragments, then incubated in Southern denaturing solution for 40 min. A sheet of 3MM chromatography paper (Whatman International Ltd), saturated in and in contact with denaturing solution, was placed on a glass plate. The gel was then placed upside down on the paper, and Saran wrap was placed around the gel edges to prevent evaporation. Hybond-N+ nylon transfer membrane (Amersham Biosciences) was placed on the gel, followed by three sheets of water-soaked 3MM paper. A stack of paper towels, followed by a glass plate and a 1-kg weight were placed on top, to allow transfer of DNA to the membrane by capillary action. Transfer proceeded overnight; the membrane was then soaked in Southern neutralising solution for 30 min, and in 2x SSC buffer (Invitrogen) for 10 min, before being air dried and baked at 80°C for 2 h.

Probe labelling. cDNA probes were labelled with $^{32}$P-dCTP by random priming, using the Megaprime labelling kit (Amersham Biosciences) according to manufacturer’s instructions. Labelled probes were purified through a ProbeQuant™ G-50 Micro Column (Amersham Biosciences), by centrifugation at 700g for 2 min. The purified probe was denatured at 95°C for 5 min and then cooled on ice.
Hybridisation. Blotted membranes were incubated in 20 ml hybridisation buffer by rotation at 42°C for at least four hours. The labelled probe was then added, and hybridisation proceeded at 42°C overnight.

Washing and detection. Hybridised membranes were washed with 3x SSC, 0.1% SDS for 15 min at 65°C with shaking. The wash was repeated with 2x SSC, 0.1% SDS. Damp membranes were wrapped in Saran wrap and exposed to X-OMAT imaging film (Kodak) at -70°C. After 2-7 days exposure, films were developed using an X150 X-ograph developer (X-ograph Imaging Systems).

Stripping. Residual probe was stripped from hybridised membranes by addition of boiling 0.5% SDS and incubation at room temperature for 3 h with shaking. Membranes were exposed to film to detect efficiency of stripping.

2.6.2 Northern blot detection

A Human Multiple Tissue Northern (MTN®) Blot (CLONTECH Laboratories, Inc) was used for analysis of VPS13A mRNA distribution. Radioactive cDNA probes were prepared and the blot was hybridised as above. The hybridised membrane was washed twice in 0.5x SSC, 0.5% SDS for 30 min (once at room temperature, once at 65°C) with shaking. The washed membrane was exposed, developed and stripped as above, except that exposure time was shorter (2 h for β-actin probe, 2 days for VPS13A probe).
Chapter 2: Materials and Methods

2.7 DNA sequencing analysis

During assembly of plasmid constructs, inserts were fully sequenced by fluorescent dye terminator sequencing prior to further manipulation. At every subsequent cloning step, ligation junctions were sequenced to check integrity. PCR products were also sequenced for mutation detection (section 2.8). Earlier experiments used slab gel electrophoresis on the ABI PRISM 377 DNA sequencer (section 2.7.1); this was later superseded by capillary electrophoresis on the ABI PRISM 3700 DNA Analyser (section 2.7.2) (both instruments from Applied Biosystems). Fluorescent samples were protected from the light as far as possible.

2.7.1 Sequencing on ABI 377

Sequence reactions were performed using the BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems). The 20-μl reaction contained 30-90 ng purified PCR product or 200-500 ng plasmid DNA, 4 μl terminator premix, 10 pmol primer and 4 μl sequencing buffer. Thermocycling was as follows: 96°C for 5 min; 30 cycles of 96°C for 10 s, 50°C for 5 s, and 60°C for 4 min. Unincorporated nucleotides were removed by sodium acetate-ethanol precipitation as follows. 0.1 volumes of 3 M sodium acetate, pH 4.6 and 2.5 volumes of cold absolute ethanol were added to each sequencing reaction, which were then incubated on ice for 25 min. The samples were pelleted by centrifugation at 2200g for 30 min at 4°C, and pellets were washed by centrifugation for 5 min in room temperature 70% ethanol. Pellets were air dried before being resuspended in 2 μl ABI 377 loading buffer. Samples were denatured at 95°C for 3 min and cooled on ice, before being subjected to electrophoresis on a 4.2% polyacrylamide sequencing gel in 1x Tris-borate EDTA
Chapter 2: Materials and Methods

(TBE) buffer. Sequences were analysed by means of the Sequencing Analysis and Sequence Navigator programs (both from Applied Biosystems).

2.7.2  **Sequencing on ABI 3700**

The 20-μl sequencing reaction contained 10-40 ng purified PCR product or 150-300 ng plasmid DNA, 4 μl BigDye version 3.1 terminator premix, 10 pmol primer and 2 μl BigDye buffer (Applied Biosystems). Thermocycling was as above (section 2.7.1). To aid purification, 0.1 volumes of 125 mM EDTA were added to samples prior to addition of sodium acetate/ethanol as before (section 2.7.1). Incubation at room temperature for 15 min was followed by centrifugation as before. Sample plates were inverted and briefly spun at 185g to remove the supernatant. The pellet wash was performed at 1650g for 15 min at 4°C. Air dried pellets were resuspended in 20 μl 50% (v/v) Hi-Di™ formamide (Applied Biosystems) and denatured as above. Denatured samples were stored at 4°C before electrophoresis. Sequences were analysed as above.
Chapter 2: Materials and Methods

2.8 Mutation detection protocols

2.8.1 Strategy

PCR products were generated from gDNA as in section 2.2 and subjected to electrophoresis to check for large insertions/deletions. XK exons and flanking intronic sequences were screened for mutations by direct sequencing of PCR products, as in section 2.7. XK primers used solely in sequencing reactions are listed in Appendix 1. Most exons of VPS13A were analysed by DHPLC, as in section 2.8.2. However, for 11 exons (exons 2, 16, 29, 33, 35, 48, 49, 50, 66, 67 and 72), analysis by DHPLC was not appropriate, either because PCR product yield was low or because the amplicon contained a high frequency polymorphism. These exons were therefore sequenced directly. The presence of each mutation was verified in the proband and any available family members by restriction enzyme analysis (section 2.10.1) or sequencing. The DNA of 50 control individuals was analysed to check for the presence of any missense mutations identified in Wave I (section 3.1). Wave II missense or exonic splice-site mutations (section 3.2) were checked in 192 control individuals (European Collection of Cell Cultures) by restriction enzyme analysis. Where the mutations did not provide a restriction fragment polymorphism, a diagnostic restriction site was introduced by PCR mutagenesis (primer sequences in Appendix 1).

2.8.2 DHPLC analysis

PCR amplification was performed as in section 2.2.2. Amplification products from probands were combined in a 3:1 ratio with the appropriate wild-type homozygous amplicon. To optimise the formation of heteroduplexes, samples were heated to 95°C for 4 min, followed by 42 cycles of 95°C with −1.6°C per cycle for 1 min. The
reannealed samples were stored at -20°C before DHPLC analysis on the WAVE™ DNA Fragment Analysis System (Transgenomic). Samples that showed a variant DHPLC pattern were further analysed by direct sequencing (section 2.7).

2.9 Haplotype analysis

Polymorphic microsatellite markers GATA89A11, D9S1674, GATA89CO8, D9S153, D9S1780 and D9S1867 flanking VPS13A were used in order to define haplotypes in the CHAC critical region. Fluorescently labelled primers specific for the markers were used in standard PCR amplification (section 2.2.1), except that 0.8 pmol/μl each primer, 1.28 mM dNTPs and 0.04 U/μl AmpliTaq Gold DNA polymerase were used, in a total reaction volume of 25 μl. PCR product yield was estimated for each marker following agarose gel electrophoresis (section 2.3), and products were diluted in dH₂O and pooled as required. 1.5 μl of dilutions were combined with 0.1 μl GeneScan® 400HD ROX size standard (Applied Biosystems) and 10 μl Hi-Di Formamide, then subjected to capillary electrophoresis on the ABI PRISM 3770. Samples were analysed using the Genotyper program (Applied Biosystems).
2.10 Molecular cloning

Most techniques in this section are based on standard protocols (Sambrook et al. 1989).

2.10.1 Restriction digestion

Digests were performed using restriction endonucleases (New England BioLabs, Inc), according to manufacturer’s instructions. Enzymes were subsequently heat inactivated at the recommended temperature for 25 min. To verify construct identity in plasmid DNA minipreps (section 2.1.3), 1 µl of miniprep was typically digested in a 30 µl reaction with 2.5-10 U enzyme for 2 h. For mutation confirmation in patient and control PCR amplicons (section 2.8.1), 5 µl of purified PCR product was typically digested with 5-20 U enzyme in a 30 µl reaction overnight. Preparative digests, used in plasmid construct assembly, typically digested 1-5 µg DNA overnight in 50-100 µl reaction volume. Digests were subjected to electrophoresis (section 2.3) and the desired DNA fragment was subsequently excised and purified. If gel purification was not desired, a second digest cleaving only the unwanted fragment was performed for 2 h, prior to column purification (section 2.5).
2.10.2 Modification of restriction fragment ends

DNA fragment ends were occasionally modified after restriction digestion (section 2.10.1) in order to permit correct DNA ligation (section 2.10.3). Samples were subsequently column- or gel purified (section 2.5) before further processing.

Filling-in of 5’ overhangs - 44 µl digestion reaction was combined with 5 µl of dNTPs (330 µM) and 1 µl of DNA polymerase I Klenow fragment (Amersham Biosciences, 1 U/µl) and incubated at room temperature for 30 min.

Polishing of 3’ overhangs - T4 DNA polymerase was used for this process, since its 3’-5’ exonuclease activity is ~1000 times more efficient than that of Klenow fragment. 45 µl digestion reaction was combined with 2.5 µl of dNTPs (2 mM), 1.5 µl dH2O and 1 µl of T4 DNA polymerase (Invitrogen, 1-8 U/µl) and incubated at 37°C for 15 min. The reaction was stopped by addition of 2.5 µl 0.5 M EDTA.

2.10.3 Ligation

Ligation were typically calculated to combine a 3:1 molar ratio of insert to vector, using the following calculation:

\[
\frac{Av \times Si}{Sv} \times R = Ai
\]

where: \( Av \) = amount of vector (ng); \( Si \) = size of insert (kb); \( Sv \) = size of vector (kb); \( R \) = molar ratio of insert to vector; \( Ai \) = amount of insert (ng).

Standard ligations - Restriction-digested (section 2.10.1) vector and insert were purified (section 2.5) and combined in a 10 µl reaction volume containing 1 x ligase buffer and 2.5 U of T4 DNA ligase (both in Rapid DNA Ligation kit, Roche). The
reaction was incubated at room temperature for ≥1 h, and then transformed into competent *E. coli* (section 2.10.4).

**Cloning of PCR fragments** - To provide a ready source of product for downstream cloning procedures, PCR products were cloned using the pGEM®-T Vector System (Promega), according to manufacturer’s instructions. Briefly, 2 μl of PCR product/pGEM-T Vector ligation reaction were heat-shock transformed into competent *E. coli* (section 2.10.4) and recovered in 1 ml SOC medium (Invitrogen). 100 μl of the transformation culture were plated on to LB-ampicillin agar supplemented with 0.5 mM isopropylthio-β-D-galactoside (IPTG) and 80 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and incubated overnight at 37°C. Interruption of the β-galactosidase gene by the PCR product insert allows discrimination between positive (white, β-galactosidase inactive) and negative (blue, β-galactosidase active) colonies.

2.10.4 Transformation of *E. coli*

Due to the high transformation efficiency of the technique, electroporation was used to introduce 1 μl of yeast plasmid miniprep (section 2.21.7) into electrocompetent DH10B cells (Invitrogen), according to manufacturer’s instructions. All other plasmids were introduced into bacteria via heat-shock transformation, as detailed below.
Heat-shock transformation (standard) - 50 μl of chemically competent XL-1 Blue Supercompetent cells (Stratagene) or DH5α cells were combined with 5 μl ligation reaction (section 2.10.3) and incubated on ice for 20 min. Cells were heat-shocked at 42°C for 30 s (DH5α) or 45 s (XL-1 Blue) and then placed on ice for 2 min. The cells were supplemented with 250-450 μl SOC medium, before incubation with shaking at 37°C for 1 h. Aliquots of the culture (1/20 and 1/2 total volume) were subsequently plated on to LB-agar plates containing the appropriate antibiotic and incubated overnight at 37°C.

Quick transformation protocol - For propagation of intact plasmids conferring ampicillin resistance, 5 μl of DH5α cells were combined with 50 ng of miniprep DNA and placed on ice for 5 min. Cells were heat-shocked, cooled on ice and supplemented with SOC medium as above. The ampicillin-resistance gene is rapidly expressed in transformant cells and so no 37°C recovery period is required when transforming intact plasmids. The entire transformant mixture was therefore directly plated on to an LB-ampicillin agar plate and incubated as above.

Transformation of bacterial expression vectors - One Shot® BL21 Star™ (DE3) chemically competent cells (Invitrogen) were used for efficient, high-level expression of recombinant fusion protein, required for production of chorein antigen (section 2.12). They were transformed with 50 ng of intact plasmid DNA according to manufacturer’s recommendations. BL21 cells are not recommended for long-term propagation of constructs, so expression vectors were maintained in XL1-Blue cells.
2.10.5 Preparation and revival of glycerol stocks

For long-term storage of plasmid constructs, glycerol stocks of transformant bacteria were prepared. An aliquot of overnight culture (810 µl) was combined with 190 µl sterile 80% glycerol, snap frozen on dry ice and stored at -80°C. Cells were revived by scraping off ~20µl of the frozen mixture and streaking on to an LB-antibiotic agar plate, which was then incubated at 37°C overnight.
<table>
<thead>
<tr>
<th>Reagent</th>
<th>Composition</th>
<th>Notes*</th>
</tr>
</thead>
<tbody>
<tr>
<td>TE buffer, pH x</td>
<td>10 mM Tris.Cl, pH x 1 mm EDTA, pH 8.0</td>
<td>Autoclave</td>
</tr>
<tr>
<td>Transport buffer</td>
<td>100 mM NaCl 10 mM Tris.Cl, pH 8.0 25 mM EDTA, pH 8.0 0.5% SDS</td>
<td>Add SDS after autoclaving</td>
</tr>
<tr>
<td>Terrific Broth (100 ml)</td>
<td>90 ml Terrific Broth premix 10 ml 10x K salts</td>
<td>Combine under sterile conditions</td>
</tr>
<tr>
<td>Terrific Broth premix (500 ml)</td>
<td>6 g bactotryptone 12 ml yeast extract 2 ml glycerol</td>
<td>Autoclave</td>
</tr>
<tr>
<td>10x K salts (100 ml)</td>
<td>2.31 g KH₂PO₄ 12.54 g K₂HPO₄</td>
<td>Autoclave</td>
</tr>
<tr>
<td>Deaza-dGTP-dNTP mix</td>
<td>2.5 mM 7-deaza dGTP 2.5 mM dGTP 5 mM dATP 5 mM TTP 5 mM dCTP</td>
<td>Use 2 µl in 50 µl PCR</td>
</tr>
<tr>
<td>10x agarose loading buffer</td>
<td>1 mg/ml bromophenol blue 1 mg/ml xylene cyanole FF 30% (v/v) glycerol 1 x TBE</td>
<td>Store at 4°C</td>
</tr>
<tr>
<td>Southern denaturing solution</td>
<td>1.5 M NaCl 0.5 M NaOH</td>
<td></td>
</tr>
<tr>
<td>Southern neutralising solution</td>
<td>1.5 M NaCl 0.5 M Tris.Cl, pH 7.5</td>
<td></td>
</tr>
<tr>
<td>Hybridisation buffer</td>
<td>50% formamide</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50 mM Na₂PO₄, pH 7.2 1 mM EDTA 8% dextran sulphate 1% SDS</td>
<td>Mix first four reagents at 50°C until dissolved. Add dextran &amp; SDS and stir until dissolved. Cool to 37°C, then add Denhart’s, DNA and tRNA. Store at 37°C.</td>
</tr>
<tr>
<td>10x Denhart’s solution (Sigma)</td>
<td>100 µg/ml herring sperm DNA 25 µg/ml yeast tRNA</td>
<td></td>
</tr>
<tr>
<td>ABI377 loading buffer (5 ml)</td>
<td>1 ml 50 mg/ml blue dextran in 25 mM EDTA 4 ml formamide</td>
<td>Store at 4°C</td>
</tr>
<tr>
<td>Sequencing gel (for 30 gels)</td>
<td>324 g urea 93.6 ml 40% acrylamide 19:1 9 g amberlite resin (Applied Biosciences)</td>
<td>Mix to dissolve urea and deionise mixture. Filter and add 180 ml 5x TBE. Degas for 15 min. Store at 4°C.</td>
</tr>
<tr>
<td>LB-agar (1 l)</td>
<td>20 g LB base (Sigma) 15 g agar (Becton Dickinson)</td>
<td>Autoclave, then cool to &lt;50°C before supplementing with 100 µg/ml ampicillin, 50µg/ml kanamycin, 10µg/ml gentamicin, or 12.5 µg/ml chloramphenicol</td>
</tr>
</tbody>
</table>

* Reagents are stored at room temperature unless otherwise stated.
2.11 CELL CULTURE TECHNIQUES

2.11.1 Cell lines

The following human cell lines were used in this study: HeLa (cervical carcinoma, (Scherer and Hoogasian 1954)); MRC-5 SV2 (SV40- transformed foetal lung fibroblast, (Huschtscha and Holliday 1983)); 293T (embryonic kidney fibroblast, transformed with Ad5 DNA, (Graham et al. 1977)); H4 (neuroglioma, (Arnstein et al. 1974; Day and Ziolkowski 1979)); RD (rhabdomyosarcoma, (McAllister et al. 1969)); Hep3B (hepatocarcinoma, (Aden et al. 1979; Knowles et al. 1980)); and K562 (multipotential myelogenous leukaemia, (Lozzo et al. 1981)). Patient and control peripheral blood leucocytes were used to prepare EBV-transformed lymphoblastoid cell lines (European Collection of Cell Cultures). COS-7 (SV40-transformed African green monkey kidney fibroblasts, (Gluzman 1981)) and CHO-KI (Chinese hamster ovary cells, (Puck et al. 1958; Kao and Puck 1968)) cell lines were also used. Collaborators (Maria Teresa Dotti, University of Siena, Italy and Saidi Mohiddin, National Heart Lung and Blood Institute, US) kindly provided primary skin fibroblasts from ChAc patients; Natalie Wilson kindly provided some from a healthy control. All cell culture reagents were purchased from Sigma, unless otherwise stated.

2.11.2 Maintenance of cells

Non-adherent cells (lymphoblastoid and K562 cell lines) were grown in RPMI-1640 media supplemented with 10% foetal bovine serum, 100 U/ml penicillin/streptomycin and 2 mM L-glutamine. All other cells were grown in Dulbecco's modified Eagle medium, supplemented as above. Cells were grown within 25-, 75- or 175-cm² tissue culture flasks (BD Biosciences) and incubated at 37°C with 5% CO₂.
2.11.3 *Harvesting and subculture of adherent cells*

When cells were confluent, the medium was removed and phosphate-buffered saline (PBS) was used to wash the cells. Washed cells were lifted from the flask by incubation with Trypsin-versene EDTA for 3 min, followed by brisk agitation. Trypsinised cells were transferred to centrifuge tubes, supplemented with growth medium (which contains a trypsin inhibitor) and subjected to centrifugation at 150g for 5 min. The cell pellet was resuspended in an appropriate volume of medium, and cells were counted using a haemocytometer if required for downstream applications (e.g., immunofluorescence (section 2.20) or transfection (section 2.11.6)). A 5 to 10-fold dilution was used to subculture cells.

2.11.4 *Harvesting and subculture of non-adherent cells*

Cells grown to maximal density in 50 ml growth medium in a 75-cm² flask were transferred to a centrifuge tube and pelleted as above. When required, cells were washed by resuspension in PBS and recentrifugation. A 10-fold dilution was used to subculture cells.

2.11.5 *Preparation and revival of frozen stocks*

Pelleted cells from a 75-cm² flask were gently resuspended in 4 ml cold cell freeze mix (RPMI-1640 supplemented with penicillin/streptomycin, 30% foetal bovine serum and 10% dimethyl sulphonyde) and 1-ml aliquots quickly transferred to cryovials. These were wrapped in bubble wrap to prevent rapid freezing, and the package was placed at -80°C overnight. Cryovials were subsequently transferred to liquid nitrogen storage tanks. Cells were revived from frozen stocks by a 1-min
incubation in a 37°C waterbath, followed by resuspension in growth medium. Non-adherent cells were pelleted by centrifugation and resuspended in 10 ml fresh growth medium before transferral to a 25-cm² flask. Adherent cells were transferred to a 75- or 175-cm² flask and incubated overnight. The medium was replaced the following day.

2.11.6 Transfection of 293T and MRC5-SV2 cells

DNA constructs were transfected into cells by means of GeneJuice™ Transfection Reagent (Novagen, Inc), according to manufacturer’s instructions. GeneJuice is a proprietary formulation of polyamines optimised for efficient transport of DNA. For co-immunoprecipitation experiments, 10-cm diameter tissue culture plates were each seeded with 1.5 x 10⁶ cells and allowed to settle for 4-16 h. Equimolar amounts of plasmid (corresponding to 5 µg of plasmid pCMVT-TAB2, 5 µg of pcD4-chac2 and 10 µg of pcDFchac8) were transfected in various combinations: each transfection mixture contained 30 µl GeneJuice. For immunofluorescence experiments (section 2.20), a poly-L-lysine-coated 13-mm diameter coverslip was placed in each well of a 24-well tissue culture plate and seeded with 5 x 10⁴ cells. Per well, 4.5 µl GeneJuice was combined with 1 µg of pcDFchac8 and/or 0.5 µg of pCMV-TAB2 and transfected as above. Cells were processed 36-48 h post-transfection.
2.12 Preparation of recombinant antigen

2.12.1 Expression of recombinant protein in bacterial cultures
A colony of BL21-Star (DE3) transformant bacteria (section 2.10.4) was inoculated into 5-10 ml LB-ampicillin and incubated with shaking at 37°C overnight. This starter culture was then diluted 1/10 in LB-ampicillin and incubated as before until bacterial growth reached log phase (measured as $OD_{600}$ of 0.6 - 0.8), usually after ~90 min. Aliquots of the culture were either supplemented with 0.1 mM IPTG to induce protein expression, or left without supplementation. Induction was allowed to proceed at a variety of temperatures and lengths of time to determine optimal protein expression conditions, as described in sections 4.2.1-4.2.3. Aliquots were removed at various time points - 100 µl of whole culture were stored at 4°C until required for standard PAGE (section 2.15.1), while 400 µl were fractionated as below. For large-scale preparation of recombinant protein, starter cultures were scaled up to 100 ml.

2.12.2 Fractionation of bacterial cell lysates
Aliquots of a log phase culture (section 2.12.1) were spun at 2200g for 10 min and cell pellets were resuspended in 50 µl pellet solubilisation buffer. Cell suspensions were sonicated in an FB11002 sonicating waterbath (Fisher Scientific) for 5 min, and then subjected to centrifugation as before. The supernatant (soluble fraction) was transferred to a new tube; the pellet (insoluble fraction) was resuspended in 50 µl pellet solubilisation buffer. Both fractions were stored at 4°C until analysis by standard PAGE (section 2.15.1) and Coomassie staining (section 2.16).
2.12.3 Solubilisation of glutathione-S-transferase (GST)-chorl

Solubilisation of recombinant fusion protein by addition of the sodium salt of the alkyl anionic detergent N-laurylsarcosine (sarkosyl) was performed after the method of Frangioni and Neel (1993). A 500-ml log phase bacterial culture was pelleted by centrifugation at 3000g for 10 min at 4°C. The pellet was washed by resuspension in 25 ml ice-cold PBS and centrifugation as before. The washed pellet was resuspended in 12 ml ice-cold sodium-Tris-EDTA (STE) and digested with 100 μg/ml of lysozyme for 15 min on ice. To enhance subsequent GST binding to glutathione (section 2.12.4), DTT was added to 10 mM; 10% sarkosyl solution in STE was then added to 1.5% final concentration. After a 1-min sonication in an ice bath, the mixture containing solubilised fusion protein was subjected to centrifugation in a 40-ml round-bottomed tube at 27000g, 4°C for 40 min. The supernatant was passed through a 0.45-μm syringe filter and transferred to a 50-ml centrifuge tube, where 10% Triton X-100 and STE were added to give a final concentration of 2% Triton X-100 in 25 ml total volume. This mixture was rocked for 30 min at room temperature to allow sequestration of sarkosyl by Triton X-100, which facilitates efficient GST-glutathione binding. For large-scale solubilisation of GST-chorl, sixteen 500-ml cultures were processed as above, pooling samples prior to addition of Triton X-100 to give eight 50-ml mixtures.

2.12.4 Purification of GST-chorl

Recombinant GST-chorl fusion protein was purified on glutathione Sepharose beads (Amersham Biosciences) according to a protocol adapted from the manufacturer and Frangioni and Neel (1993).
Small-scale purification - As it proved impossible to purify soluble GST-chorl in the absence of detergent, multiple small-scale purifications were performed containing various combinations of different detergents and other additives, as detailed in section 4.2.3. A typical small-scale optimisation trial used 1-ml aliquots of Triton-treated GST-chorl mixture (section 2.12.3) in microcentrifuge columns with 50-μl bed volume of glutathione-agarose (MicroSpin™ GST Purification Module, Amersham Biosciences), after column equilibration according to the manufacturer’s recommendations.

Large-scale purification - Once the purification conditions had been optimised, the following scaled-up protocol was used. Four 1.33-ml aliquots of 75% glutathione bead suspension were each transferred to 15-ml centrifuge tubes, and washed by centrifugation in 10 ml STE at 650g, 4°C for 2 min. This was followed by a wash with STE supplemented with sarkosyl and DTT (STE-SD). The equilibrated beads were resuspended in 1 ml STE-SD to give a 50% slurry. The eight 50-ml Triton-treated GST-chorl mixtures (section 2.12.3) were each rocked with 1 ml bead slurry for 1 h at room temperature to allow association of GST-chorl with the glutathione beads. Pairs of tubes containing bead-protein mixture were pooled and passed through an empty PD-10 column (Amersham Biosciences), thus generating four fusion protein-bound columns, each with a 1-ml bed volume. Each column was washed with 400 ml STE-SD, and the washed beads from each column were resuspended in 1 ml STE-SD and pooled to a 15-ml tube. A thrombin recognition site located between GST and chorl allowed release of the chorein moiety from the glutathione beads - 160 U thrombin protease (Amersham Biosciences) were added and the suspension was
rocked overnight at room temperature. The following day, the suspension was transferred to a new PD-10 column and the 4-ml eluate was collected (eluate 1). Three more eluates were collected (eluate 2, 4 ml; eluate 3, 2 ml; eluate 4, 2 ml) by allowing the appropriate volume of STE-SD to pass though the column. The beads were finally resuspended in 4 ml STE-SD and aliquots of the eluates and beads were subjected to electrophoresis to assess purity and estimate yield of chor1 antigen by comparison with bovine serum albumin standards, as shown in Figure 4.3c.

2.12.5 Antigen concentration and immunisation

Eluates 1 and 2 were pooled and concentrated in a Vivaspin 6 centrifugal concentrator column with a 10-kDa exclusion filter (Vivascience AG), according to manufacturer’s recommendations. The 0.5-ml sample was transferred to a microfuge tube; the column was then washed out with 0.5 ml STE-SD, which was added to the sample. The total protein concentration of the purified chor1 antigen was measured using the bicinchonic acid method (BCA protein assay kit, Pierce), according to manufacturer’s instructions. Several 40-μl aliquots (~160 μg protein) were made and stored at -80°C until required. Immunisation of rabbits SK274 and SK275 with chor1 was performed by Eurogentec Bel SA (Herstal, Belgium); the serum collected 87 days post-immunisation was designated anti-chor1.
2.13 Protein sample preparation

2.13.1 Yeast cell lysate preparation

Yeast cultures were grown to log phase \( \text{OD}_{600} \approx 0.4-0.5 \) and 5 ml were pelleted by centrifugation at 970g for 5 min. Cell pellets were washed by resuspension in 1.5 ml dH\(_2\)O and centrifugation at 6000g for 30 s. Washed pellets were resuspended in 2 volumes of ice-cold yeast cell lysis buffer. To shear the cells, 2 pellet volumes of sterile acid-washed 710-1180-\(\mu\)m diameter glass beads were added and the mixture was vortexed for three bursts of 30 s, and then sonicated in an ice bath for 5 min. Samples were denatured for 5 min at 100\(^\circ\)C and spun at 16000g for 1 min. The supernatant was transferred to a new tube and subjected to standard PAGE (section 2.15.1) and Western blotting (section 2.17).

2.13.2 Mammalian cell lysate preparation

Cells were harvested and pelleted as in sections 2.11.3 - 2.11.4. Cell pellets were washed with ice-cold PBS then lysed in 10 pellet volumes of ice-cold mammalian cell lysis buffer. After 20 min incubation on ice, lysates were clarified by centrifugation at 16000g at 4\(^\circ\)C for 20 min. The supernatant (soluble fraction) was transferred to a new tube. If required, the pellet (insoluble fraction) was resuspended in 100 \(\mu\)l lysis buffer. Protein concentrations were measured using the BCA assay as in section 2.12.5. The \(\text{OD}_{562}\) of the samples was measured on a SpectraMAX 190 plate reader (Molecular Devices). Fractions were stored at -20\(^\circ\)C until required.
2.13.3 Erythrocyte membrane preparation

Erythrocyte membranes were prepared after the method of Dodge et al. (1963). However, to enable batch processing of patient blood, the protocol was modified to utilise blood samples frozen at -20°C. Whole blood was thawed at room temperature and 2.5 ml were added to 10 ml of ice-cold erythrocyte wash solution in a 15 ml centrifuge tube. The mixture was spun at 3400g, 4°C for 10 min and all but the bottom 2 ml was removed by pipetting. The remaining mixture was split into two 1.5 ml microcentrifuge tubes and spun at 16000g, 4°C for 5 min. All but 250 µl of the supernatant was removed, taking care not to disturb the loose pellet. Most erythrocytes were already lysed by the freeze-thawing, but any intact cells were now lysed by trituration in 1 ml of a hypotonic erythrocyte lysis solution. This was followed by centrifugation at 16000g, 4°C for 5 minutes. The loose pellet of membranes was washed multiple times in 1 ml erythrocyte lysis solution followed by centrifugation as above, until it could be clearly distinguished from a solid pellet of clotted material. The two loose pellets from each original sample were pooled and transferred to a fresh tube. Washing in erythrocyte lysis solution proceeded as above, until the supernatant remained colourless. Membranes were stored at -20°C until required. Membrane yield was presumably dependent on initial erythrocyte count and varied from 100-200 µl: 2.5 µl of this were combined with 4 µl dH₂O and prepared for electrophoresis (section 2.15) and Western blotting (section 2.17).
2.14 Dot blotting

Three dilutions of purified chor1 antigen (section 2.12.5; 25, 50, 100 ng/μl) were prepared and 2 μl of each were spotted on to Hybond ECL Nitrocellulose Membrane (Amersham Biosciences). The membrane was dried at 60°C for 5 min and then cooled, prior to immunodetection as in section 2.17.2.

2.15 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Standard PAGE was initially used for all applications. However, we discovered that even on 5% polyacrylamide gels full-length chorein could not be detected, and so gradient PAGE was employed whenever this was necessary. Rainbow™ coloured protein molecular weight markers (Amersham Biosciences) were used as size standard. After SDS-PAGE, proteins were either stained with Coomassie blue (section 2.16) or detected by Western blot analysis (section 2.17).

2.15.1 Standard PAGE

Protein samples were mixed with PAGE sample buffer and denatured at 100°C for 10 min. The denatured samples were subjected to electrophoresis in 1x PAGE running buffer on 10% or 12.5% polyacrylamide gels in the XCell Surelock™ Mini-Cell system (Invitrogen). To improve separation of proteins, gels were cast with a ~5 mm top layer of stacking gel. Electrophoresis proceeded at 80 V for 50 min, then at 120 V until the dye front had passed through the gel.
2.15.2 Gradient PAGE

All products were purchased from Invitrogen. Protein samples were mixed with NuPAGE™ LDS Sample Buffer and Sample Reducing Agent and incubated at 70°C for 10 min. Samples were then subjected to electrophoresis on precast NuPAGE Novex Tris-acetate gels (3-8% polyacrylamide gradient), under denaturing and reducing conditions. Electrophoresis proceeded at 150 V for 80-90 min.

2.16 Coomassie staining

Polyacrylamide gels were gently agitated in Coomassie blue stain for 1 h, then agitated in multiple changes of Coomassie destain until protein bands were clearly visible. Gels were placed on 3MM Whatman paper and dried in a 583 Gel Dryer (Bio-Rad Laboratories) for 40 min at 80°C.

2.17 Western blot analysis

2.17.1 Transfer of proteins

The XCell II™ Blot Module (Invitrogen) was used for electrophoretic transfer of proteins. Proteins on standard polyacrylamide gels were transferred for 1 h at 26 V to nitrocellulose membranes, in 1x Western transfer buffer. Proteins on gradient gels were transferred according to manufacturer’s instructions, on to Invitrolon™ polyvinylidene difluoride membranes (Invitrogen). To check equal loading of erythrocyte membrane preparations (section 2.13.3), blots were agitated for 15 min in 0.1% Ponceau S solution (Sigma), followed by brief destaining in dH₂O until protein
bands were clearly visible. Blots were subsequently destained completely in 1x Tris-buffered saline with Tween-20 (TBS-T) before downstream processing.

2.17.2 Immunodetection

All incubation steps were performed with agitation to enable even distribution of reagent. To reduce background signal, blots were blocked overnight at 4°C in 5% powdered non-fat milk dissolved in 1x TBS-T. The following day, blots were incubated in primary antibody diluted in 1% milk in TBS-T for 1 h at room temperature. Primary antibodies used were: rabbit anti-chor1 (section 2.12.5, 1:5000 dilution); rabbit anti-myc (Invitrogen, 1:500); rabbit anti-FLAG (Sigma, 1:2000); mouse anti-early endosome antigen 1 (Transduction Laboratories, 1:5000); mouse anti-myc (ICRF, 1:2000); mouse anti-T7 (Novagen, Inc, 1:10000); and mouse anti-GAL4 DB (Santa Cruz Biotechnology, Inc, 1:200). Blots were subsequently washed by three 20-min incubations in TBS-T, and then incubated with horseradish peroxidase-conjugated goat anti-mouse or anti-rabbit secondary antibodies (1/5000 dilution, Bio-Rad Laboratories), as for the primary incubation. Blots were washed as before, then proteins were visualised using the Enhanced Chemiluminescent (ECL) plus Western Blotting Detection System (Amersham), according to manufacturer’s instructions. If required for reprobing, blots were stripped according to the ECL instructions and stored at 4°C in TBS-T until required.
2.18 Antibody purification and processing

Anti-chorl antibodies were purified from antiserum according to the following protocol. To avoid degradation of antibodies, all reagents and procedures were at 4°C unless otherwise stated.

2.18.1 Column preparation

NHS-activated Sepharose 4 Fast Flow bead suspension (Amersham Biosciences, 1 ml) was briefly spun at 620g. The supernatant was removed by decanting and beads were washed twice by brief centrifugation in 10 ml 5 mM HCl. The acidified beads were tumbled overnight in a 2-ml microcentrifuge tube with ~1.5 mg purified chorl antigen (section 2.12.5). The following day, beads were spun for 5 min at 700g, transferred to a 15-ml tube, washed twice by brief centrifugation in 10 ml STE-S and incubated on ice for 2 h in 100 mM Tris.Cl, pH 8.0. The resin was then washed sequentially: in 10 ml of 100 mM Tris.Cl, 500 mM NaCl; then 10 ml of 100 mM sodium acetate, pH 4.0, 500 mM NaCl. This sequential wash was repeated three times and followed by two washes in PBS: the beads were then transferred to a PD-10 column and kept in PBS until required. For long-term storage of the resin, 0.1% Na azide and 20% ethanol in PBS were used.

2.18.2 Serum treatment

10 ml anti-chorl serum (section 2.12.5) was dialysed against five changes of 1 l PBS over a period of 24 h. The dialysed serum was passed through a 0.2 μm syringe filter and tumbled with the drained antigen-conjugated beads overnight. The column was drained the following day, and 3-ml washes of PBS were passed through until no
excess protein was found in the wash, as determined by OD$_{280}$ measurement (~5-6 washes). Column-bound antibodies were eluted with 2-ml fractions of 100 mM glycine, pH 3.0 and immediately neutralised with 50 µl 1 M Tris, pH 9.0. This was repeated until little protein was found in the eluate, measured as above (6 elutions). Eluates were pooled as appropriate after Coomassie analysis (section 2.16), supplemented with 0.1% NaN$_3$ and adjusted to pH 7.4, before aliquoting and storage at -20°C.

2.18.3 Antibody depletion assay

For antibody depletion experiments, the chorl antigen-conjugated Sepharose column used in antibody purification (section 2.18.1) was employed. Anti-chorl antiserum was diluted 1/2500 in TBS. Half was passed three times through the column, while half was left untreated. Both serum preparations were subsequently adjusted to the appropriate dilution for use in Western blotting (section 2.17).
2.19 Co-immunoprecipitation

Per immunoprecipitation (IP), 6 μg (6 μl MαT7) of primary antibody were combined with 10 μl Protein G Sepharose™ 4 Fast Flow beads (Amersham Biosciences) and a total of 250 μl IP buffer in a microcentrifuge tube and tumbled for 2 h at room temperature, to allow formation of the antibody-bead complex. The antibody-bead complex was washed five times by a 5-s centrifugation at 16000g in 500 μl IP buffer and kept on ice until required. Cell lysates were prepared as in section 2.13.2 (but using IP buffer for cell lysis) and tumbled with 20 μl protein G at 4°C for 1 h. This mixture was then subjected to centrifugation at 16000g for 5 min to pellet unspecific protein G-binding proteins. The supernatant was combined with washed antibody-bead complex and tumbled at 4°C overnight to allow precipitation of antibody-bound protein. The beads were then washed five times in 500 μl IP buffer and resuspended in PAGE sample buffer. Samples were denatured and loaded on to duplicate acrylamide gels for subsequent electrophoresis (section 2.15) and Western blotting (section 2.17).
MRC5-SV2 and 293T cells grown on 13-mm coverslips in a 24-well plate (section 2.11.6) were processed for indirect immunofluorescence as follows. Cells were washed with 1 ml PBS (prewarmed to 37°C to avoid dislodging of loosely-adhered 293T cells) and then inverted to remove the wash. Cells were fixed with 0.5 ml of room-temperature 4% paraformaldehyde in PBS, or 0.5 ml of ice-cold methanol. Paraformaldehyde fixation retains the native structure of the proteins, while methanol denatures them, thus enabling presentation of different epitopes for antibody recognition. Fixation proceeded at room temperature for 10 min, followed by four washes in 0.6 ml IF wash solution. To reduce background signal, fixed cells were incubated in IF block solution for 30 min at room temperature, prior to incubation for 1 h in primary antibody. Primary antibodies used were rabbit anti-chor1 (section 2.12.5; 1:1000 dilution in 200 μl IF block), rabbit anti-FLAG (1:500) and mouse anti-T7 (1:500). Cells were washed five times with IF wash and then incubated with fluorescent dye-conjugated secondary antibody as before, but kept in the dark. Secondary antibodies used were Alexa Fluor® 594-conjugated goat anti-rabbit (red) and Alexa Fluor 488-conjugated goat anti-mouse (green) (both from Molecular Probes, 1:500 dilution). Cells were washed as before, then coverslips were removed from the wash and mounted on to a microscope slide with ~10 μl Fluoromount-G (Southern Biotechnology Associates, Inc). Slides were stored at 4°C, before examination within 24 h under the MRC 1024 confocal microscope system using Lasersharp 2000 software (both from Bio-Rad Laboratories).
## Chapter 2: Materials and Methods

### Table 2.2 - Composition of reagents used for protein techniques

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Composition</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pellet solubilisation buffer</td>
<td>1x PBS</td>
<td>Store at 4°C</td>
</tr>
<tr>
<td></td>
<td>1% Triton X-100</td>
<td>Add 1x protease inhibitor cocktail (Sigma) and 1 mM Phenyl methyl sulphonyl fluoride (PMSF) immediately before use</td>
</tr>
<tr>
<td>STE</td>
<td>10 mM Tris.Cl, pH 8.0</td>
<td>Store at 4°C</td>
</tr>
<tr>
<td></td>
<td>1 mM EDTA</td>
<td>Add 1x protease inhibitor cocktail immediately before use</td>
</tr>
<tr>
<td></td>
<td>150 mM NaCl</td>
<td></td>
</tr>
<tr>
<td>STE-SD</td>
<td>1x STE</td>
<td>Sterile filter</td>
</tr>
<tr>
<td></td>
<td>0.1% sarkosyl</td>
<td>Make fresh each time</td>
</tr>
<tr>
<td></td>
<td>1 mM DTT</td>
<td></td>
</tr>
<tr>
<td>Yeast cell lysis buffer</td>
<td>2.5x PAGE sample buffer</td>
<td>Add 1 mM PMSF immediately before use</td>
</tr>
<tr>
<td>Mammalian cell lysis buffer</td>
<td>50 mM Tris.Cl, pH 8.0</td>
<td>Store at 4°C</td>
</tr>
<tr>
<td></td>
<td>150 mM NaCl</td>
<td>Add 1x protease inhibitor cocktail immediately before use</td>
</tr>
<tr>
<td></td>
<td>1% Triton X-100</td>
<td></td>
</tr>
<tr>
<td>Erythrocyte wash solution</td>
<td>5 mM Na₂HPO₄, pH 8.0</td>
<td>Store at 4°C</td>
</tr>
<tr>
<td></td>
<td>0.9 % NaCl</td>
<td>Add 1x protease inhibitor cocktail immediately before use</td>
</tr>
<tr>
<td>Erythrocyte lysis solution</td>
<td>5 mM Na₂HPO₄, pH 8.0</td>
<td>Store at 4°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Add 1x protease inhibitor cocktail immediately before use</td>
</tr>
<tr>
<td>5x PAGE sample buffer</td>
<td>10% (w/v) SDS</td>
<td>Freeze in aliquots at -20°C for long term storage</td>
</tr>
<tr>
<td></td>
<td>50% (v/v) glycerol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25% (v/v) β-mercaptoethanol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>250 mM Tris.Cl, pH 6.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1 mg/ml bromophenol blue</td>
<td></td>
</tr>
<tr>
<td>10x PAGE running buffer</td>
<td>250 mM Tris</td>
<td>Don’t adjust pH</td>
</tr>
<tr>
<td></td>
<td>2 M glycine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1% (w/v) SDS</td>
<td></td>
</tr>
<tr>
<td>x% polyacrylamide gel</td>
<td>X% 37.5:1 acrylamide</td>
<td>Make fresh each time</td>
</tr>
<tr>
<td></td>
<td>1 x separating buffer</td>
<td>Add 150 µl 10% (w/v) ammonium persulphate and 15 µl TEMED to polymerise 16 ml gel</td>
</tr>
<tr>
<td>4x separating buffer</td>
<td>1.5 M Tris</td>
<td>Adjust pH to 8.8</td>
</tr>
<tr>
<td></td>
<td>0.4% (w/v) SDS</td>
<td></td>
</tr>
<tr>
<td>Stacking gel</td>
<td>4% 37.5:1 acrylamide</td>
<td>Make fresh each time</td>
</tr>
<tr>
<td></td>
<td>1x stacking buffer</td>
<td>Add 100 µl 10% (w/v) ammonium persulphate and 15 µl TEMED to polymerise 10 ml gel</td>
</tr>
<tr>
<td>2x stacking buffer</td>
<td>0.5 M Tris</td>
<td>Adjust pH to 6.8</td>
</tr>
<tr>
<td></td>
<td>0.2% (w/v) SDS</td>
<td></td>
</tr>
<tr>
<td>Coomassie blue stain</td>
<td>2.5 mg/ml Coomassie Brilliant Blue</td>
<td>Filter before use</td>
</tr>
<tr>
<td></td>
<td>45% (v/v) ethanol</td>
<td>Protect from light</td>
</tr>
<tr>
<td></td>
<td>10% (v/v) acetic acid</td>
<td></td>
</tr>
<tr>
<td>Coomassie destain</td>
<td>10% (v/v) ethanol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10% (v/v) acetic acid</td>
<td></td>
</tr>
<tr>
<td>10x Western transfer buffer</td>
<td>240 mM Tris</td>
<td>Don’t adjust pH</td>
</tr>
<tr>
<td></td>
<td>200 mM glycine</td>
<td>Add 20% methanol when preparing 1x stock</td>
</tr>
</tbody>
</table>

70
<table>
<thead>
<tr>
<th>Reagent</th>
<th>Composition</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBS-T</td>
<td>1x TBS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1% Tween-20 (BDH)</td>
<td></td>
</tr>
<tr>
<td>10x TBS</td>
<td>9% (w/v) NaCl</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100 mM Tris Cl, pH 7.5</td>
<td></td>
</tr>
<tr>
<td>10x IP buffer</td>
<td>0.5 M Tris Cl, pH 7.8</td>
<td>Autoclave, store at 4°C</td>
</tr>
<tr>
<td></td>
<td>1.5 M NaCl</td>
<td>Add 0.2% (w/v) iodoacetamide, 1 mM PMSF and 1x</td>
</tr>
<tr>
<td></td>
<td>10 mM EDTA, pH 8.0</td>
<td>protease inhibitor cocktail to 1x stock</td>
</tr>
<tr>
<td></td>
<td>1% Triton X-100</td>
<td>immediately before use</td>
</tr>
<tr>
<td>IF wash solution</td>
<td>1x TBS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1% (v/v) Triton X-100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.02% sodium azide</td>
<td></td>
</tr>
<tr>
<td>IF block solution</td>
<td>1x TBS</td>
<td>Store at 4°C</td>
</tr>
<tr>
<td></td>
<td>0.1% (v/v) Triton X-100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 mg/ml fish gelatine</td>
<td></td>
</tr>
</tbody>
</table>

* Reagents are stored at room temperature unless otherwise stated.
2.21 **YEAST TWO-HYBRID PROTOCOLS**

The yeast two-hybrid (Y2H) system used in this study was the ProQuest™ Two-Hybrid System with Gateway™ Technology (Invitrogen). This system employs yeast strain MaV203. This strain contains auxotrophic mutations in the \textit{LEU2} and \textit{TRP1} genes to allow selection of bait and prey constructs, respectively. Three reporter genes, \textit{HIS3}, \textit{lacZ} and \textit{URA3} are also engineered with GAL4-activated promoter sequences in this strain. Induction of \textit{HIS3} and \textit{URA3} allow two-hybrid-dependent transcription activation to be monitored by cell growth on plates lacking histidine (-His) or uracil (-Ura), respectively. Induction of \textit{lacZ} results in a blue colour when assayed with X-Gal. Induction of \textit{URA3} also results in conversion of 5-fluoroorotic acid (5FOA) to the toxic compound 5-fluorouracil: growth of cells containing interacting proteins is thus inhibited on 5FOA-supplemented medium. This phenotype is capable of detecting weaker bait-prey interactions than assaying for growth on -Ura medium can detect.

2.21.1 **Y2H library screen strategy**

\textit{VPS13A} cDNA fragments were transferred through site-specific recombination (section 2.21.3) into the bait vector pDEST32. Small-scale transformation (section 2.21.5) was used to introduce bait constructs into MaV203. DNA from a human brain cDNA library (section 2.21.4) was introduced into bait-transformed yeast by large-scale transformation (section 2.21.6). Double transformants were assayed for reporter gene activity as recommended in the ProQuest instruction manual (Invitrogen). To account for false positives arising from mutation of the host yeast, clones were further analysed by retransformation (section 2.21.2).
2.21.2 Retransformation strategy

As recombinant DNA extracted from yeast is refractory to molecular analysis, plasmids were extracted from putative positive colonies by miniprep (section 2.21.7), and electroporated into bacteria (section 2.10.4). Prey constructs were obtained after miniprep of transformant bacteria (section 2.1.3), were cotransformed with the relevant bait into fresh yeast by small-scale transformation (section 2.21.5) and were then assayed for reporter gene induction as before.

2.21.3 Assembly of Y2H bait and prey constructs

The strategy for bait and prey construction is illustrated in Figures 5.1 & 5.2. The Gateway™ Technology uses the lambda site-specific recombination system to facilitate transfer of heterologous DNA sequences flanked by att sites. BP Clonase™ enzyme mix was used to catalyse recombination of attB-flanked PCR product with attP-containing donor vector pDONR221 to generate entry vector containing attL-flanked insert. Following verification of the entry vector by sequencing (section 2.7), LR Clonase™ enzyme mix was used to catalyse its recombination with attR-containing yeast two-hybrid destination vectors pDEST32 and pDEST22. Recombination reactions were transformed into Library Efficiency DH5α competent E. coli (Invitrogen) by heat shock (section 2.10.4). Donor and destination vectors contain the ccdB gene, which is lost upon recombination. CcdB protein inhibits DNA gyrase and thus growth of sensitive E. coli strains, which allows negative selection of unrecombined vectors.
2.21.4 Preparation of Y2H cDNA library

The library used in this study was Human Brain ProQuest Pre-made cDNA Library (Invitrogen), which uses cDNA isolated from a 27-year old male. This library is cloned into pEXP-AD502 prey vector and has an average cDNA insert size of 1.9 kb. Glycerol stocks of DH10B bacteria transformed with the library (~1.3 colony-forming units) were inoculated into 50 ml Terrific Broth and incubated with shaking at 30°C overnight. DNA was subsequently extracted by maxiprep (section 2.1.4).

2.21.5 Small-scale polyethylene glycol/lithium acetate (PEG/LiAc) transformation of yeast

The following technique provides enough competent yeast for seven transformations. Five colonies were suspended in 50 µl dH₂O, plated out on to YPAD-agar and incubated at 30°C overnight. The following day, the lawn of cells was resuspended in 5 ml YPAD and the OD₆₀₀ of this suspension was measured. Enough suspension was used to inoculate 100 ml YPAD to an initial OD₆₀₀ of 0.2. This starter culture was incubated with shaking at 30°C until yeast reached log phase (OD₆₀₀ = 0.4-0.5, usually after ~3 h incubation). Yeast cells were harvested by centrifugation at 970g for 5 min and the cell pellet was washed by resuspension in 40 ml sterile dH₂O and centrifugation at 240g for 10 min. Cells were then washed as before in 20 ml TE/LiAc solution and the cell pellet was adjusted to a final volume of 350 µl with TE/LiAc. The competent cells were stored at room temperature and used within 30 min of preparation. Carrier DNA was prepared by denaturing 10 mg/ml sonicated salmon sperm DNA (Invitrogen) at 100°C for 5 min and cooling on ice. Transformations were performed by combining 50 µl competent cell suspension with 5 µl carrier DNA and 100 ng each plasmid in a screw-cap microfuge tube. PEG/LiAc solution (300 µl) was
added and the sample was inverted to mix. Samples were incubated in a 30°C waterbath for 30 min with occasional mixing, then heat-shock transformed by a 5-min incubation in a 42°C waterbath. Samples were cooled on ice for 2 min and then spun at 6000g for 30 s. The pellet of transformed cells was resuspended in 500 µl sterile dH₂O. The suspension was then plated on to SC agar with the appropriate drop-out supplements. For cotransformation of yeast with two constructs, 100 µl of suspension were plated; for transformation with single constructs, 100 µl of a 1/10 dilution were plated. Plates were incubated for 48 h at 30°C before analysis of colonies.

2.21.6 Large-scale PEG/LiAc transformation of yeast

Large-scale transformation of yeast proceeded largely as for small-scale transformation (section 2.21.5), with the following modifications. Three lawn plates of yeast transformed with the relevant bait construct were prepared. Lawns were resuspended in 15 ml SC-Leu and used to inoculate a 350-ml SC-Leu starter culture in a 2-l capacity conical flask. The culture was grown to log phase and cells were pelleted, and then washed in 140 ml sterile dH₂O, followed by 70 ml TE/LiAc. Competent cells were resuspended to a final volume of 1.6 ml; as a negative control, a 75-µl aliquot was taken into a screw-cap microfuge tube and combined with 5 µl carrier DNA and 775 ng empty prey vector pEXP-AD502, then 450 µl PEG/LiAc. The remainder was combined in a 50-ml tube with 100 µl carrier DNA and 15 µg human brain cDNA library, then 9 ml PEG/LiAc. After incubation at 30°C, cells were heat-shocked at 42°C for 15 min with occasional mixing. After cooling, library-transformed cells were pelleted by centrifugation at 240g for 10 min, then resuspended to a final volume of 24 ml; the control transformed-cells were spun at 6000g for 30 s and resuspended to a final volume of 1.2 ml. To estimate the total
number of transformants, 100-μl aliquots of 1:10, 1:100 and 1:1000 dilutions of the
cell suspension were spread on to 10-cm SC-Leu-Trp agar plates. The undiluted
suspension was spread on to 15-cm SC-Leu-Trp-His plates (400 μl per plate). SC-
Leu-Trp-His agar had been supplemented with 50 mM 3-amino-1, 2, 4-triazole to
suppress leaky expression of the HIS3 gene. Plates were incubated for 60 h at 30°C
before removal of visible cell growth by replica cleaning with sterile velvets and
reincubation for 60-72 h.

2.21.7 Extraction of recombinant DNA from yeast

Five yeast transformant colonies were inoculated into 5 ml SC-Leu-Trp medium and
incubated with shaking at 30°C for ~24 h. 1 ml was then pelleted by centrifugation at
14000g for 30 s. Cells were lysed by suspension in 100 μl yeast alkaline lysis
solution. Lysis was stopped after 15 min by sequential addition of 500 μl 1x TE, pH
7.5 and 60 μl 3 M sodium acetate. To remove proteins, 600 μl TE-saturated phenol:
chloroform: isoamyl alcohol (25:24:1) was added and the mixture was vortexed for 2
min, and subjected to centrifugation at 14000g for 2 min. The upper aqueous phase
was transferred to a new microfuge tube and treated again with phenol: chloroform:
isoamyl alcohol as before. Ice-cold isopropanol (650 μl) was then added, and the
mixture was incubated at -20°C for ≥ 30 min. After centrifugation at 14000g for 5
min, every trace of supernatant was removed and the DNA pellet was washed by
centrifugation as before in 100 μl 70% ethanol. The air-dried pellet was resuspended
in 10 μl TE, pH 7.5 and stored at -20°C.
### Table 2.3 - Composition of reagents used for yeast two-hybrid techniques

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Composition</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC medium (1 l)</td>
<td>6.7 g yeast nitrogen base without aa 1.35 g aa-purine mix</td>
<td>Autoclave, add 50 ml autoclaved 40% glucose 8 ml 20 mM uracil 8 ml 100 mM His 8 ml 100 mM Leu 8 ml 40 mM Trp (protect from light)</td>
</tr>
<tr>
<td>SC-Leu-Trp</td>
<td>As above</td>
<td>As above, but supplement with glucose, uracil and His only</td>
</tr>
<tr>
<td>aa-purine mix</td>
<td>2 g each of all standard aa except His, Leu, Trp 2 g adenine hemisulphate</td>
<td></td>
</tr>
<tr>
<td>TE/LiAc solution</td>
<td>1x TE, pH 7.5 100 mM lithium acetate</td>
<td>Prepare from filter-sterilised stocks of 1 M lithium acetate and 10x TE Prepare TE/LiAc solution immediately before use</td>
</tr>
<tr>
<td>PEG/LiAc solution</td>
<td>1x TE, pH 7.5 100 mM lithium acetate 40% PEG-3350</td>
<td>Prepare from filter-sterilised stocks of 1 M lithium acetate, 10x TE and 50% PEG-3350 Prepare PEG/LiAc solution immediately before use</td>
</tr>
<tr>
<td>Yeast alkaline lysis solution</td>
<td>3% SDS 0.2 N NaOH</td>
<td>Prepare fresh each time</td>
</tr>
</tbody>
</table>

* Reagents are all stored at room temperature.
Chapter 3: Mutational analysis of \textit{VPS13A} in patients with chorea-acanthocytosis

The entire coding and flanking intronic sequence in \textit{VPS13A} was screened for mutations in 83 unrelated patients with chorea-acanthocytosis. These patients were organised into three cohorts, Waves I, II and III.

3.1 Wave I

Our initial cohort comprised the 11 previously reported ChAc families linked to chromosome 9q21-q22 (Rubio \textit{et al.} 1997). We screened the proband from each family by direct sequencing of all 73 exons of \textit{VPS13A}. We identified four homozygous and 12 heterozygous mutations, finding at least one mutation in all probands analysed (Table 3.1 & Figure 3.1). Co-segregation of the mutations with the disease haplotypes was confirmed in all available family members using the verification method outlined in Table 3.1.

3.1.1 Description of mutations

Patients in families CHAC3, 5 & 7 are heterozygous for nonsense mutations 622C>T, 6700C>T and 9109C>T, respectively (Table 3.1 & Figure 3.1). All three mutations are substitutions in the arginine codon CGA, generating a TGA premature termination codon (PTC).
Nine different insertions or deletions leading to a shift in the reading frame were identified (Table 3.1 & Figure 3.1). Patients in families CHAC1 and 3 are heterozygous for the single base-pair insertions 6404_6405insT and 4419_4420insA, respectively. CHAC4 patients are homozygous for insertion of a dinucleotide (3557_3558insAC) and a heterozygous 4 base-pair insertion 9289_9290insTTTG was found in CHAC7 patients. CHAC2 and CHAC8 patients are homozygous and CHAC11 patients are heterozygous for deletions of a single base pair at positions 1592, 1115 and 237, respectively. Heterozygous deletions of a dinucleotide (2833_2834del) and a tetranucleotide (9429_9432del) were found in affected members of families CHAC10 and 11, respectively. All nine frameshift mutations are predicted to lead to introduction of a PTC in the VPS13A transcript.

Affected members of CHAC5 harbour a heterozygous substitution of A to C at the splice-acceptor dinucleotide AG of intron 48 (Table 3.1 & Figure 3.1). As RNA was not available from this family, it was not possible to determine the precise effect of this mutation on splicing. However, input of this base change into the SpliceView program (http://125.itba.mi.cnr.it/~webgene/wwwspliceview.html) revealed that the new junction sequence is too divergent from the consensus to score as a splice-site. We therefore predict that this mutation is likely to cause aberrant splicing of the VPS13A mRNA transcript.
Three different missense mutations were identified in this screen (Table 3.1 & Figure 3.1). CHAC1 patients are heterozygous for the base change 269T>A, which results in substitution of isoleucine to lysine at position 90. CHAC6 affected individuals harbour the homozygous change 4354T>C, which leads to substitution of proline for serine at position 1452. Individuals with ChAc in family CHAC9 are heterozygous for the base change 8162A>G, which leads to replacement of tyrosine by cysteine at position 2721. None of the three amino-acid substitutions was found in 100 control chromosomes.

Table 3.1 - VPS13A disease mutations identified in Wave I

<table>
<thead>
<tr>
<th>Family*</th>
<th>Location</th>
<th>DNA change*</th>
<th>Protein change*</th>
<th>Mutation type</th>
<th>Verification method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAC1</td>
<td>Exon 4</td>
<td>269T&gt;A (ht)</td>
<td>I90K*</td>
<td>Missense</td>
<td>Sequencing</td>
</tr>
<tr>
<td></td>
<td>Exon 48</td>
<td>6404_6405insT (ht)</td>
<td>S2136KfsX1</td>
<td>Frameshift</td>
<td>Ddel</td>
</tr>
<tr>
<td>CHAC2*</td>
<td>Exon 17</td>
<td>1592del (hm)</td>
<td>I531KfsX6</td>
<td>Frameshift</td>
<td>Sequencing</td>
</tr>
<tr>
<td>CHAC3</td>
<td>Exon 9</td>
<td>622C&gt;T (ht)</td>
<td>R208X</td>
<td>Non sense</td>
<td>Sequencing</td>
</tr>
<tr>
<td></td>
<td>Exon 38</td>
<td>4419_4420insA (ht)</td>
<td>G1474RfsX6</td>
<td>Frameshift</td>
<td>Sequencing</td>
</tr>
<tr>
<td>CHAC4*</td>
<td>Exon 33</td>
<td>3557_3558insAC (hm)</td>
<td>V1187LfsX11</td>
<td>Frameshift</td>
<td>Sequencing</td>
</tr>
<tr>
<td></td>
<td>Exon 48</td>
<td>6700C&gt;T (ht)</td>
<td>R2234X</td>
<td>Nonsense</td>
<td>Taqf</td>
</tr>
<tr>
<td></td>
<td>Intron 48</td>
<td>6775-2A&gt;C (ht)</td>
<td>(SA: no score (wt = 82))</td>
<td>Splice site</td>
<td>Styf</td>
</tr>
<tr>
<td>CHAC6*</td>
<td>Exon 37</td>
<td>4354T&gt;C (hm)</td>
<td>S1452P*</td>
<td>Missense</td>
<td>Sequencing</td>
</tr>
<tr>
<td>CHAC7</td>
<td>Exon 68</td>
<td>9109C&gt;T (ht)</td>
<td>R3037X</td>
<td>Non sense</td>
<td>Sequencing</td>
</tr>
<tr>
<td></td>
<td>Exon 71</td>
<td>9289_9290insTTTGTG (ht)</td>
<td>T3098CfsX11</td>
<td>Frameshift</td>
<td>Sequencing</td>
</tr>
<tr>
<td>CHAC8*</td>
<td>Exon 13</td>
<td>1115del (hm)</td>
<td>K372SfsX1</td>
<td>Frameshift</td>
<td>Sequencing</td>
</tr>
<tr>
<td>CHAC9</td>
<td>Exon 59</td>
<td>8162A&gt;G (ht)</td>
<td>Y2721C*</td>
<td>Missense</td>
<td>Sequencing</td>
</tr>
<tr>
<td>CHAC10</td>
<td>Exon 27</td>
<td>2833_2834del (ht)</td>
<td>K945EfsX10</td>
<td>Frameshift</td>
<td>Sequencing</td>
</tr>
<tr>
<td>CHAC11</td>
<td>Exon 4</td>
<td>237del (ht)</td>
<td>E80KfsX10</td>
<td>Frameshift</td>
<td>Sequencing</td>
</tr>
<tr>
<td></td>
<td>Exon 72</td>
<td>9429_9432del (ht)</td>
<td>R3143SfsX4</td>
<td>Frameshift</td>
<td>Sequencing</td>
</tr>
</tbody>
</table>

* Families are asterisked where we had prior evidence of homozygosity by descent (Rubio et al., 1997). * Families are numbered according to the cDNA sequence of VPS13A isoform 1A reported by Rampoldi et al. (2001) as CHAC isoform A (GenBank accession no. NM_033305), with the adenosine of the initiation codon assigned position 1. Mutation nomenclature is as recommended by the Human Genome Variation Society (http://www.genomic.unimelb.edu.au/mgi/mutnomen/index.html). * Not found in 100 control chromosomes. * Wild type and mutant DNA sequence flanking the splice site was entered into the Splice View program (http://125.itba.mi.cnr.it/~webgene/wwwspliceview.html) - the scores produced are shown. Key: ?, second heterozygous mutation not found; ex, exon; int, intron; SA, splice acceptor.
Figure 3.1 - Identification of VPS13A mutations in patients with ChAc. The pedigrees of families CHAC3, CHAC5 and CHAC6 are shown alongside sequences generated by direct sequencing of genomic DNA amplicons from (left) a control individual and (right) the relevant proband. The sequences illustrate respectively nonsense, frameshift, splice-site and missense mutations found in these families. Affected individuals are indicated in the pedigree by filled symbols, unaffected by open symbols. Arrowheads indicate the position of the mutation in the proband sequence. For heterozygous mutations, the mutant allele sequence is displayed below the wild-type sequence. Key: ht = heterozygous, hm = homozygous
3.2 Wave II

3.2.1 XK screening

Forty-five probands with symptoms compatible with ChAc were initially included in the follow-up to Wave I. As most families in this cohort were not big enough to permit linkage analysis, and the symptoms of MLS can be similar to ChAc (Danek et al. 2004), we wished to exclude involvement of the XK gene before proceeding with the VPS13A screen. For 11 probands, Kell serotyping had been performed: these patients did not show the weak Kell antigen response characteristic of MLS. For the remaining 34 probands, no Kell serotyping data was available. Of these, 17 were either female or had an affected female in their pedigree, which was very unlikely to be consistent with X-linked recessive inheritance of the McLeod phenotype. This left 17 probands, who were screened for mutations in the XK gene by direct sequencing of all three exons and flanking intronic sequence. This revealed that one proband was hemizygous for a nonsense mutation in XK exon 3, 707G>A (W236X). A second patient was hemizygous for deletion of nucleotides 856-860 in the same exon. This 5-bp deletion causes a shift of reading frame at aa position 285, leading to a PTC 15 codons downstream.1 As both these mutations were likely to disrupt XK function, these patients were excluded, leaving 43 probands with suspected ChAc.

---

1 We subsequently reported these patients as MLS cases 11 and 22, respectively (Danek et al. 2001).
3.2.2  VPS13A screening

The entire coding and flanking intronic sequence in VPS13A was screened for mutations by denaturing high-performance liquid chromatography (DHPLC) in 43 unrelated patients with chorea-acanthocytosis. This revealed 57 different mutations likely to cause disease in 39 probands (Table 3.2). Only three of these mutations had been previously identified, in different families, in our Wave I cohort. All available family members were screened for the relevant mutation(s) using the verification method outlined in Table 3.2; in each case, the mutation(s) co-segregated consistently with the affection status. In seven patients only one heterozygous mutation was found; in four patients with typical symptoms of ChAc, no disease mutations were found.

3.2.3  Nonsense mutations

Seventeen different base substitutions that result in PTCs were identified (Table 3.2). Three of these mutations (1549G>T, 3109A>T, 5920G>T) result in substitution to a TAA stop codon; three other mutations (3157C>T, 6419C>G, 9219C>G) result in substitution to a TAG stop codon. The remaining 11 mutations in this group result in substitution to a TGA stop codon: the majority of these (7/11, 64%) arise in CGA arginine codons. The 5920G>T mutation was found to be homozygous not only in proband II-5, but also his affected mother (see Figure 3.2a).
Figure 3.2 - Analysis of VPS13A mutations in pseudodominant ChAc families. a, Restriction fragment length polymorphism (RFLP) analysis of exon 45 in Wave II family 5. The mutation 5920G>T (E1974X) creates a DraI site: bands corresponding to the mutant allele are indicated in red. b, Segregation of mutations in Wave II family 8. The father in this pedigree showed some mild dyskinesia, hence the grey shading. The left-hand panel shows RFLP analysis of exon 72. The splicing mutation 9474G>A destroys a ScrFI site: the band corresponding to the mutant allele is indicated in red. The right-hand panel shows sequencing analysis of exon 29. An arrowhead indicates the position of the mutation (3109A>T (K1037X)).
3.2.4 Splice-site mutations

Fifteen different mutations are predicted to result in altered *VPS13A* mRNA splicing (Table 3.2). Twelve of these are located in introns: they directly alter the highly conserved AG or GT dinucleotides of the intronic splice-acceptor or splice-donor sites, respectively. Mutation 883-1_892del spans intron 11 and exon 12. Although it deletes 10 nucleotides of exon 12 and could therefore shift the reading frame of the *VPS13A* mRNA, we consider its primary effect to be on splicing, since the new intron/exon junction formed (attctttagTATTTTCAGTATTAT → attcttttaTTAT) is widely divergent to the consensus splice-acceptor sequence. The remaining two splice-site mutations (7806G>A & 9474G>A) alter the 3’ nucleotide in exons 55 and 72, respectively. The first substitution occurs within the sequence CCGgttaata → CCAgttaata: although both trinucleotides CCG and CCA code for proline, this exon/intron junction no longer conforms to the splice-donor consensus. The same is true of mutation 9474G>A in proband II-8: the substitution AGGgttaaat → AGAgtaaat does not affect the coding for arginine 3158, but does abolish the splice-donor site.

Interestingly, the father of proband II-8 is homozygous for this mutation (see Figure 3.2b) and shows mild features of the disease. Both exonic splice-site mutations were found to be absent in 192 control individuals using the verification methods indicated in Table 3.2. In each case, input of these base changes into the SpliceView program revealed that the new junction sequence was too divergent from the consensus to score as a splice-site (Table 3.2).
3.2.5 Missense mutations

Only two different base changes resulting in potentially pathogenic, non-conservative amino-acid substitutions were identified (3283G>C & 7378T>C, Table 3.2). Mutation 3283C>G results in the substitution of alanine to proline at position 1095. Mutation 7378T>C results in the substitution of tryptophan to arginine at position 2460. Neither of these missense mutations was found when screening 192 normal individuals using the verification methods outlined in Table 3.2.

3.2.6 Insertion/deletion mutations

Twenty different insertions or deletions were identified that lead to a shift in the reading frame, introducing a PTC (Table 3.2). These comprise 11 1-bp deletions, two 1-bp insertions, two 2-bp deletions, three 4-bp deletions and one deletion each of five and 14 base pairs. Mutation 2029_2031delins27 involves the deletion of the trinucleotide CAT and the insertion of the 27-nucleotide sequence ATATACTAATATCTGCTTCTTTTGGAC. This mutation results in a net increase of 24 nucleotides, so does not shift the reading frame; however, the third trinucleotide that is introduced is a PTC. In probands II-6 and II-42, it was not possible to amplify exon 23 and exons 70-73, respectively. Southern analysis of the appropriate regions revealed that proband 6 was homozygous for an approximately 7-kb deletion covering exon 23, and that proband 42 was homozygous for a deletion of at least 13 kb removing the four terminal exons of VPS13A (Figure 3.3a & b).
### Table 3.2 - VPS13A disease mutations identified in Wave II

<table>
<thead>
<tr>
<th>Location</th>
<th>DNA change</th>
<th>Protein change</th>
<th>Proband no.</th>
<th>Verification method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nonsense</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 9</td>
<td>622C&gt;T&lt;sup&gt;a&lt;/sup&gt;</td>
<td>R208X&lt;sup&gt;b&lt;/sup&gt;</td>
<td>II-12 (ht), II-27* (hm), II-39 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 17</td>
<td>1549G&gt;T</td>
<td>E517X</td>
<td>II-3 (ht)</td>
<td><em>Apol</em></td>
</tr>
<tr>
<td>Exon 18</td>
<td>1616C&gt;G</td>
<td>S539X</td>
<td>II-18 (ht)</td>
<td><em>Tfil</em></td>
</tr>
<tr>
<td>Exon 25</td>
<td>2593C&gt;T</td>
<td>R865X</td>
<td>II-9 (hm)</td>
<td><em>TaqI</em></td>
</tr>
<tr>
<td>Exon 29</td>
<td>3109A&gt;T</td>
<td>K1037X</td>
<td>II-8 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 30</td>
<td>3157C&gt;T</td>
<td>Q1051X</td>
<td>II-15* (hm)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 34</td>
<td>3889C&gt;T</td>
<td>R1297X</td>
<td>II-38 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 37</td>
<td>4355C&gt;G</td>
<td>S1452X</td>
<td>II-35 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 37</td>
<td>4411C&gt;T</td>
<td>R1471X</td>
<td>II-26 (ht)</td>
<td><em>StyI</em></td>
</tr>
<tr>
<td>Exon 45</td>
<td>5920G&gt;T</td>
<td>E1974X</td>
<td>II-5* (hm)</td>
<td><em>DraI</em></td>
</tr>
<tr>
<td>Exon 46</td>
<td>6094C&gt;T</td>
<td>R2032X</td>
<td>II-4 (ht)</td>
<td><em>BsrI</em></td>
</tr>
<tr>
<td>Exon 48</td>
<td>6419C&gt;G</td>
<td>S2140X</td>
<td>II-24 (ht)</td>
<td><em>BamHI</em></td>
</tr>
<tr>
<td>Exon 48</td>
<td>6494G&gt;A</td>
<td>W2165X</td>
<td>II-10 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 50</td>
<td>7005G&gt;A</td>
<td>W2335X</td>
<td>II-36 (ht)</td>
<td><em>Pcl</em></td>
</tr>
<tr>
<td>Exon 56</td>
<td>7867C&gt;T</td>
<td>R2623X</td>
<td>II-23 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 68</td>
<td>9109C&gt;T&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R3037X&lt;sup&gt;b&lt;/sup&gt;</td>
<td>II-1 (ht)</td>
<td><em>Tsp5I</em></td>
</tr>
<tr>
<td>Exon 70</td>
<td>9219C&gt;G</td>
<td>Y3073X</td>
<td>II-13 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td><strong>Splice site</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 11</td>
<td>883-1_892del</td>
<td>(SA: no score (wt = 82))&lt;sup&gt;c&lt;/sup&gt;</td>
<td>II-22 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 17</td>
<td>1593+1G&gt;A</td>
<td>(SD: no score (wt = 86))&lt;sup&gt;c&lt;/sup&gt;</td>
<td>II-36 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 17</td>
<td>1596-2A&gt;C</td>
<td>(SA: no score (wt = 89))</td>
<td>II-16 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 17</td>
<td>1596-1G&gt;C</td>
<td>(SA: no score (wt = 89))</td>
<td>II-35 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 21</td>
<td>2170+1G&gt;A</td>
<td>(SD: no score (wt = 85))</td>
<td>II-26 (ht)</td>
<td><em>HphI</em></td>
</tr>
<tr>
<td>Intron 22</td>
<td>2288+2T&gt;C</td>
<td>(SD: no score (wt = 83))</td>
<td>II-2 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 36</td>
<td>4242+1G&gt;T</td>
<td>(SD: no score (wt = 93))</td>
<td>II-40 (hm)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 40</td>
<td>4956+1G&gt;T</td>
<td>(SD: no score (wt = 83))</td>
<td>II-12 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 55</td>
<td>7806G&gt;A&lt;sup&gt;a&lt;/sup&gt;</td>
<td>(SD: no score (wt = 76))&lt;sup&gt;c&lt;/sup&gt;</td>
<td>II-3 (ht)</td>
<td><em>HpaII</em></td>
</tr>
<tr>
<td>Intron 61</td>
<td>8472-1G&gt;C</td>
<td>(SA: no score (wt = 88))&lt;sup&gt;c&lt;/sup&gt;</td>
<td>II-2 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 65</td>
<td>8907+2T&gt;A</td>
<td>(SD: no score (wt = 75))&lt;sup&gt;c&lt;/sup&gt;</td>
<td>II-16 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 70</td>
<td>9275+1G&gt;A</td>
<td>(SD: no score (wt = 80))&lt;sup&gt;c&lt;/sup&gt;</td>
<td>II-1 (ht)</td>
<td><em>MaeIII</em></td>
</tr>
<tr>
<td>Intron 70</td>
<td>9276-2A&gt;T</td>
<td>(SA: no score (wt = 92))&lt;sup&gt;c&lt;/sup&gt;</td>
<td>II-41 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Intron 71</td>
<td>9399+2_8del</td>
<td>(SD: no score (wt = 92))&lt;sup&gt;c&lt;/sup&gt;</td>
<td>II-14 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 72</td>
<td>9474G&gt;A&lt;sup&gt;a&lt;/sup&gt;</td>
<td>(SD: no score (wt = 81))&lt;sup&gt;c&lt;/sup&gt;</td>
<td>II-8 (ht)</td>
<td><em>ScrFI</em></td>
</tr>
<tr>
<td><strong>Missense</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 31</td>
<td>3283G&gt;C&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A1095P</td>
<td>II-21 (ht)</td>
<td><em>BstXI</em></td>
</tr>
<tr>
<td>Exon 53</td>
<td>7378T&gt;C&lt;sup&gt;c&lt;/sup&gt;</td>
<td>W2460R</td>
<td>II-28 (hm)</td>
<td><em>Acl</em></td>
</tr>
<tr>
<td><strong>Frameshift</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 13</td>
<td>994del</td>
<td>A332LfsX9</td>
<td>II-7 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 13</td>
<td>1125_1128del</td>
<td>S375RfsX22</td>
<td>II-11 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 14</td>
<td>1187_1188del</td>
<td>F396fsX5</td>
<td>II-17* (hm)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 14</td>
<td>1208_1211del</td>
<td>Q403RfsX5</td>
<td>II-23 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 20</td>
<td>2029_2031delins27</td>
<td>H677delinsLYX</td>
<td>II-29 (ht)</td>
<td><em>PstAI</em></td>
</tr>
<tr>
<td>Exon 34</td>
<td>3847del</td>
<td>L1283WfsX6</td>
<td>II-33 (hm)</td>
<td><em>BsrI</em></td>
</tr>
<tr>
<td>Exon 36</td>
<td>4216del</td>
<td>V1406CfsX19</td>
<td>II-38 (ht)</td>
<td><em>StyI</em></td>
</tr>
<tr>
<td>Exon 37</td>
<td>4346del</td>
<td>S1449fsX4</td>
<td>II-19 (ht)</td>
<td><em>MboII</em></td>
</tr>
<tr>
<td>Exon 39</td>
<td>4835del</td>
<td>P1612QfsX29</td>
<td>II-21 (ht)</td>
<td><em>BsrI</em></td>
</tr>
<tr>
<td>Exon 41</td>
<td>5253_5266del</td>
<td>F1751LfsX13</td>
<td>II-13 (ht)</td>
<td><em>DraI</em></td>
</tr>
</tbody>
</table>
Table 3.2 - *VPS13A* disease mutations identified in Wave II

<table>
<thead>
<tr>
<th>Location</th>
<th>DNA change</th>
<th>Protein change</th>
<th>Proband no.</th>
<th>Verification method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 46</td>
<td>6059del</td>
<td>P2020LfsX8</td>
<td>II-37 (hm)</td>
<td>HphI</td>
</tr>
<tr>
<td>Exon 47</td>
<td>6283del</td>
<td>S2095QfsX9</td>
<td>II-7 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 49</td>
<td>6804_6805insG</td>
<td>S2269VfsX6</td>
<td>II-39 (ht)</td>
<td>Avail</td>
</tr>
<tr>
<td>Exon 49</td>
<td>6828del</td>
<td>V2277LfsX11</td>
<td>II-41 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 53</td>
<td>7339_7340insT</td>
<td>Y2447LfsX4</td>
<td>II-20 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 57</td>
<td>7985_7989del</td>
<td>P2662RfsX5</td>
<td>II-32 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 57</td>
<td>8007del</td>
<td>K2669NfsX21</td>
<td>II-22 (ht)</td>
<td>Hpal</td>
</tr>
<tr>
<td>Exon 61</td>
<td>8390del</td>
<td>G2797DfsX1</td>
<td>II-14 (ht)</td>
<td>Mnl</td>
</tr>
<tr>
<td>Exon 70</td>
<td>9190del</td>
<td>V3064SfsX16</td>
<td>II-24 (ht)</td>
<td>Sequencing</td>
</tr>
<tr>
<td>Exon 72</td>
<td>9429_9432del</td>
<td>R3143SfsX4b</td>
<td>II-34 (ht)</td>
<td>BsrI</td>
</tr>
<tr>
<td>Exon 72</td>
<td>9431_9432del</td>
<td>E3144VfsX5</td>
<td>II-29 (ht)</td>
<td>PstI</td>
</tr>
</tbody>
</table>

**Gross deletion**

| Exon 23   | EX23del    | (unknown)      | II-6* (hm)  | Southern blot       |
| Exons 70-73| EX70_EX73del | (unknown)     | II-42 (hm)  | Southern blot       |

* Nucleotides and amino acids are numbered according to the cDNA sequence of *VPS13A* isoform 1A reported by Rampoldi et al. (2001) as *CHAC* isoform A (GenBank accession no. NM_033305), with the adenosine of the initiation codon assigned position 1. Mutation nomenclature is as recommended by the Human Genome Variation Society (http://www.genomic.unimelb.edu.au/mdi/mutnomen/index.html). b: Mutations that were identified in a Wave I family previously. c: Probands are asterisked where we had prior knowledge that they were offspring of a consanguineous pairing. d: Wild-type and mutant DNA sequence flanking each splice site was entered into the SpliceView program (http://125.itba.mi.cnr.it/~webgene/wwwspliceview.html) - the scores produced are shown. e: Not found in 384 control chromosomes. f: skipping of this exon in *VPS13A* mRNA would lead to a frameshift.

Key: hm, homozygous; ht, heterozygous; SA, splice acceptor; SD, splice donor; wt, wild type.
Figure 3.3 - Analysis of gross deletions in \textit{VPS13A}. \textbf{a}, Southern blot analysis of proband II-6 EX23del mutation. PCR primers KIAArt13 & KIAArt11 (see Appendix 1) were used to generate a 563-bp cDNA probe containing exons 22-26. A scaled restriction map of the normal locus is shown on the left; hybridisation of the probe to an EcoRI digest of the proband family is shown on the right. The abnormal restriction fragment is arrowed; restriction fragments that are absent in the proband are labelled \textit{a} & \textit{b}. \textbf{b}, Southern blot analysis of proband II-42 EX70_73del mutation. PCR primers 222006fp6 & BCD3239fp1 (see Appendix 1) were used to generate a 710-bp cDNA probe containing exons 65-68 & 70-73. Scaled restriction maps of the normal locus are shown on the left; hybridisation of the probe to \textit{HindIII} and \textit{EcoRI} digests of a control and proband II-42 is shown on the right. Restriction fragments that are absent in the proband are labelled \textit{c}-\textit{g}. \textbf{c}, Long-range PCR analysis of proband II-23 EX54del mutation. The forward primer for exon 53 and reverse primer for exon 55 (see Appendix 1) were used to amplify a region spanning exons 53-55. The wild-type allele (\textit{wt}) is 5797 bp; the deletion allele (\textit{del}) is approximately 4 kb. Key: \textit{C}, healthy control.
3.3 Wave III

3.3.1 XK screening
The Wave III cohort comprised 29 probands with suspected ChAc. As for Wave II, patients with no affected females in the pedigree and with no Kell serotyping data were screened for mutations in the \( \textit{XK} \) gene. No mutations likely to cause McLeod syndrome were identified, and so all 29 probands were included in the \( \textit{VPS13A} \) screen.

3.3.2 VPS13A screening
DHPLC was again employed to screen \( \textit{VPS13A} \) in the 29 probands of Wave III. This revealed 27 different mutations likely to cause disease in 22 probands (Table 3.3). Nine of these mutations have been identified before in Wave I or II. Co-segregation of the relevant mutation with affection status was verified by sequencing of all available family members. In three patients only one heterozygous mutation was found; in seven patients with typical symptoms of ChAc, no disease mutations were found.

3.3.3 Nonsense mutations
Six different nonsense mutations were identified (Table 3.3). Three of these substitutions result in generation of a TAA stop codon; the remaining three lead to TGA stop codons.
3.3.4 Splice-site mutations

Five different mutations are predicted to lead to altered \( VPS13A \) mRNA splicing (Table 3.3). Mutation 188-5T>G occurs within the intron 3/exon 4 junction. The unusual sequence of the wild-type junction, caaaaaaatgtagGTA, does not in fact score as a splice site when using the SpliceView program. In contrast, the new sequence scores 77 as a splice acceptor, but the predicted intron-exon junction is 4 bases upstream, i.e., caaaaaaaagtGTA. As it does not score with this program, it is not possible to predict how splicing at the wild-type junction is affected by the T>G substitution. However, the activation of a cryptic splice site near this sequence is likely to have an adverse effect. In proband III-15, the splice-donor sequence at the exon 6/intron 6 junction is altered from GATgtaagt (SpliceView score 83) to GATgtaaat (no score). The three remaining mutations directly alter the highly conserved intronic splice-acceptor (AG) or splice-donor (GT) dinucleotides.

3.3.5 Insertion/deletion mutations

Eleven different insertions or deletions leading to a shift in the reading frame were identified (Table 3.3). These include three 1-bp deletions, three 2-bp deletions, one 2-bp insertion and two 4-bp deletions. An insertion of the octanucleotide AACAAAAAA, and a compound deletion of two nucleotides with insertion of an A, was also identified. All eleven mutations generate a PTC in the \( VPS13A \) mRNA transcript. Homozygous deletions of one or more exons were found in six probands (Figure 3.3): two of these (III-6 & III-24) shared the same deletion of exons 70-73 as Wave II proband 42. The 7-exon gene encoding G alpha subunit 14 (GNA14) lies in a tail-to-tail arrangement with \( VPS13A \) (Figures 1.3 & 3.4). PCR analysis of families II-
42, III- 6 & III-24 revealed that GNA14 exons 6 & 7 could not be amplified in affected family members, implying that the deletion actually extends across the intergenic region between VPS13A and GNA14 and encompasses the last two exons of GNA14 (Figure 3.4).
### Table 3.3 - VPS13A disease mutations identified in Wave III

<table>
<thead>
<tr>
<th>Location</th>
<th>DNA change</th>
<th>Protein change</th>
<th>Proband no</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nonsense</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 13</td>
<td>1078C&gt;T</td>
<td>Q360X</td>
<td>III-22 (ht)</td>
</tr>
<tr>
<td>Exon 22</td>
<td>2191C&gt;T</td>
<td>R731X</td>
<td>III-30 (ht)</td>
</tr>
<tr>
<td>Exon 23</td>
<td>2347C&gt;T</td>
<td>Q783X</td>
<td>III-19* (hm)</td>
</tr>
<tr>
<td>Exon 37</td>
<td>4355C&gt;G</td>
<td>S1452X</td>
<td>III-29 (ht)</td>
</tr>
<tr>
<td>Exon 37</td>
<td>4411C&gt;T</td>
<td>R1471X</td>
<td>III-3 (hm)</td>
</tr>
<tr>
<td>Exon 41</td>
<td>5084T&gt;A</td>
<td>L1695X</td>
<td>III-36 (ht)</td>
</tr>
<tr>
<td><strong>Splice site</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 3</td>
<td>188-5T&gt;G</td>
<td>(SA: 77 (wt = no score))&lt;sup&gt;a&lt;/sup&gt;</td>
<td>III-26 (ht)</td>
</tr>
<tr>
<td>Intron 6</td>
<td>495+1G&gt;A</td>
<td>(SD: no score (wt = 83))</td>
<td>III-8 (ht)</td>
</tr>
<tr>
<td>Intron 6</td>
<td>495+5G&gt;A</td>
<td>(SD: no score (wt = 83))</td>
<td>III-15 (ht)</td>
</tr>
<tr>
<td>Intron 17</td>
<td>1596-1G&gt;C&lt;sup&gt;b&lt;/sup&gt;</td>
<td>(SA: no score (wt = 89))</td>
<td>III-29 (ht)</td>
</tr>
<tr>
<td>Intron 36</td>
<td>4242+1G&gt;T&lt;sup&gt;a&lt;/sup&gt;</td>
<td>(SD: no score (wt = 93))</td>
<td>III-11* (hm)</td>
</tr>
<tr>
<td><strong>Frameshift</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 5</td>
<td>372_373ins8</td>
<td>V125NfsX3</td>
<td>III-22 (ht)</td>
</tr>
<tr>
<td>Exon 33</td>
<td>3557_3558insAC&lt;sup&gt;h&lt;/sup&gt;</td>
<td>V1187LfsX11&lt;sup&gt;b&lt;/sup&gt;</td>
<td>III-31 (hm)</td>
</tr>
<tr>
<td>Exon 35</td>
<td>3995_3996delinsA</td>
<td>F1332fsX0</td>
<td>III-30 (ht)</td>
</tr>
<tr>
<td>Exon 38</td>
<td>4428_4431del</td>
<td>G1478LfsX5</td>
<td>III-36 (ht)</td>
</tr>
<tr>
<td>Exon 39</td>
<td>4724del</td>
<td>P1575LfsX2</td>
<td>III-14 (hm)</td>
</tr>
<tr>
<td>Exon 40</td>
<td>4903_4906del</td>
<td>K1635VfsX5</td>
<td>III-15 (ht)</td>
</tr>
<tr>
<td>Exon 45</td>
<td>5909_5910del</td>
<td>E1970VfsX3</td>
<td>III-7 (ht)</td>
</tr>
<tr>
<td>Exon 46</td>
<td>6059del</td>
<td>P2020LfsX8&lt;sup&gt;b&lt;/sup&gt;</td>
<td>III-2 (hm)</td>
</tr>
<tr>
<td>Exon 49</td>
<td>6828del&lt;sup&gt;b&lt;/sup&gt;</td>
<td>V2277LfsX11&lt;sup&gt;b&lt;/sup&gt;</td>
<td>III-9 (ht)</td>
</tr>
<tr>
<td>Exon 67</td>
<td>9065_9066del</td>
<td>Q3022RfsX9</td>
<td>III-9 (ht)</td>
</tr>
<tr>
<td>Exon 72</td>
<td>9431_9432del&lt;sup&gt;b&lt;/sup&gt;</td>
<td>E3144VfsX5&lt;sup&gt;b&lt;/sup&gt;</td>
<td>III-18 (hm)</td>
</tr>
<tr>
<td><strong>Gross deletion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exons 2-3</td>
<td>EX2_EX3del</td>
<td>(unknown)</td>
<td>III-20* (hm)</td>
</tr>
<tr>
<td>Exons 8-9</td>
<td>EX8_EX9del</td>
<td>(unknown)</td>
<td>III-13* (hm)</td>
</tr>
<tr>
<td>Exons 46-50</td>
<td>EX46_EX50del</td>
<td>(unknown)</td>
<td>III-33 (hm)</td>
</tr>
<tr>
<td>Exon 54</td>
<td>EX54del</td>
<td>(unknown)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>III-23 (hm)</td>
</tr>
<tr>
<td>Exons 70-73</td>
<td>EX70_EX73del&lt;sup&gt;b&lt;/sup&gt;</td>
<td>(unknown)</td>
<td>III-6 (hm), III-24* (hm)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Nucleotides and amino acids are numbered according to the cDNA sequence of VPSJ3A isoform 1A reported by Rampoldi et al (2001) as CHAC isoform A (GenBank accession no. NM_033305), with the adenosine of the initiation codon assigned position 1. Mutation nomenclature is as recommended by the Human Genome Variation Society (http://www.genomic.unimelb.edu.au/mdi/mutnomen/index.html).<sup>b</sup> Mutations that were identified in a Wave I or II family previously. Probands are asterisked where we had prior knowledge that they were offspring of a consanguineous pairing.<sup>c</sup> Wild-type and mutant DNA sequence flanking each splice site was entered into the SpliceView program (http://125.itba.mi.cnr.it/~webgene/wwwspliceview.html) - the scores produced are shown. Not found in 384 control chromosomes.<sup>f</sup> skipping of this exon in VPS13A mRNA would lead to a frameshift. Key: hm, homozygous; ht, heterozygous; SA, splice acceptor; SD, splice donor; wt, wild type.
Figure 3.4 - Further characterisation of \textit{VPS13A} EX70_73del mutation. \(a\), PCR amplification of \textit{VPS13A} exons 69 & 70, and \textit{GNA14} exons 6 & 5, in families II-42, III-24 & III-6 (see Appendix 1 for primer sequences). Pedigrees are indicated at the top of the figure - family members who were not analysed are outlined in grey. \(b\), Scaled map of chromosome 9 region covering the 3' ends of \textit{VPS13A} and \textit{GNA14} genes. Exons that cannot be amplified in probands II-42, III-6 & III-24 are shown in pink.
3.4 Exonic polymorphisms

During the course of screening Waves I-III, 24 additional exonic variants were identified (Table 3.4). Seventeen of these are silent mutations, as the nucleotide substitutions do not alter the amino acid coding. The remaining seven changes were concluded to be neutral variants, despite leading to amino-acid substitutions. This was because they were either present only in unaffected relatives, or on the same chromosome as disease-causing mutations. Interestingly, one of the so-called 'silent' mutations (1020A>G) in exon 13 occurs in the nucleotide sequence GAAGTAAAT. The new sequence scores 84 on the SpliceView program as a splice-donor site, GAG][GTAAAT (the splice-donor site in intron 13 that is usually used scores 79).

Proband II-18 is heterozygous for this change and an S539X mutation in exon 18 (Table 3.2). The exon 13 change was also found on 1 of 384 control chromosomes, so it is doubtful that this represents a pathogenic variant. However, we cannot exclude the possibility that it could lead to the ChAc phenotype in combination with a heterozygous nonsense mutation. The base change 4509A>G in exon 38, identified in proband WIII-21, may also activate a cryptic splice-site: while the wild-type sequence GTCTTTCAAGAAA does not score, the altered sequence scores 80 as a potential splice-acceptor site, GTCTTTCAG][GAAA. The splice-acceptor site in intron 37 that is usually used scores 85. Analysis of mRNA splicing in this proband would be required to discover whether this base change is a neutral or pathogenic variant.
### Table 3.4 - Exonic polymorphisms identified in VPS13A

<table>
<thead>
<tr>
<th>Exon</th>
<th>DNA change</th>
<th>Protein change</th>
<th>Frequency in ChAc proband chromosomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>1020A&gt;G</td>
<td>None</td>
<td>1/166</td>
</tr>
<tr>
<td>14</td>
<td>1197T&gt;C</td>
<td>None</td>
<td>1/166</td>
</tr>
<tr>
<td>18</td>
<td>1695C&gt;G</td>
<td>G565L</td>
<td>2/166</td>
</tr>
<tr>
<td>26</td>
<td>2693T&gt;C</td>
<td>V898A</td>
<td>4/166</td>
</tr>
<tr>
<td>26</td>
<td>2718A&gt;G</td>
<td>None</td>
<td>2/166</td>
</tr>
<tr>
<td>26</td>
<td>2814A&gt;G</td>
<td>None</td>
<td>1/166</td>
</tr>
<tr>
<td>29</td>
<td>3045C&gt;A</td>
<td>None</td>
<td>21/166</td>
</tr>
<tr>
<td>29</td>
<td>3108T&gt;C</td>
<td>None</td>
<td>2/166</td>
</tr>
<tr>
<td>33</td>
<td>3645C&gt;A</td>
<td>None</td>
<td>2/166</td>
</tr>
<tr>
<td>38</td>
<td>4469G&gt;A</td>
<td>R1490K</td>
<td>1/166</td>
</tr>
<tr>
<td>38</td>
<td>4509A&gt;G</td>
<td>None</td>
<td>1/166</td>
</tr>
<tr>
<td>39</td>
<td>4760A&gt;G</td>
<td>Y1587C</td>
<td>0/166</td>
</tr>
<tr>
<td>44</td>
<td>5535A&gt;G</td>
<td>None</td>
<td>22/166</td>
</tr>
<tr>
<td>45</td>
<td>5917G&gt;A</td>
<td>V1973I</td>
<td>9/166</td>
</tr>
<tr>
<td>45</td>
<td>5979C&gt;T</td>
<td>None</td>
<td>2/166</td>
</tr>
<tr>
<td>46</td>
<td>6027C&gt;T</td>
<td>None</td>
<td>19/164</td>
</tr>
<tr>
<td>46</td>
<td>6039C&gt;A</td>
<td>None</td>
<td>1/164</td>
</tr>
<tr>
<td>47</td>
<td>6252A&gt;G</td>
<td>None</td>
<td>1/164</td>
</tr>
<tr>
<td>48</td>
<td>6492T&gt;C</td>
<td>None</td>
<td>40/164</td>
</tr>
<tr>
<td>48</td>
<td>6531C&gt;T</td>
<td>None</td>
<td>2/164</td>
</tr>
<tr>
<td>54</td>
<td>7457T&gt;C</td>
<td>I2486T</td>
<td>1/164</td>
</tr>
<tr>
<td>63</td>
<td>8571T&gt;C</td>
<td>None</td>
<td>21/166</td>
</tr>
<tr>
<td>67</td>
<td>9069A&gt;G</td>
<td>None</td>
<td>84/166</td>
</tr>
<tr>
<td>73</td>
<td>9515C&gt;T</td>
<td>P3172L</td>
<td>2/160</td>
</tr>
</tbody>
</table>

* Nucleotides and amino acids are numbered according to the cDNA sequence of VPS13A isoform 1A reported by Rampoldi et al. (2001) as CHAC isoform A (GenBank accession no. NM_033305), with the adenosine of the initiation codon assigned position 1. 

** Possibly affects RNA splicing: see section 3.4. 

*** Mutation found in a healthy member of Wave I family CHAC6.
3.6 Summary

This chapter describes the mutational analysis of \textit{VPS13A} in 83 unrelated patients with suspected chorea-acanthocytosis. Eighty-eight different changes likely to disrupt chorein function were found in 72 patients, comprising six deletions of entire exons, 22 nonsense, 36 frameshift, 19 splice-site and five missense mutations. In 12 patients, only one heterozygous mutation was found; for 11 patients with typical symptoms of ChAc, no mutations were found. Twelve mutations were shared by one or more apparently unrelated patients, leaving 76 mutations that were only identified in a single family. In addition, 24 exonic \textit{VPS13A} variants were identified: 17 of these were synonymous, while seven were non-synonymous changes.
Chapter 4: Analysis of \textit{VPS13A} expression

4.1 Distribution of \textit{VPS13A} mRNA expression

In order to determine the tissue distribution of \textit{VPS13A} mRNA, we analysed its expression by using a human adult multiple-tissue Northern blot. As shown in Figure 4.1a, hybridisation with a cDNA probe corresponding to \textit{VPS13A} exons 7-17 shows a broad high-molecular-weight band. There are two polyadenylation signals present in exon 73, which generate a 10- or 11-kb variant of transcript 1A (exons 1-68, 70-73). Transcript 1B (exons 1-69) is predicted to be 10 kb long. The broad band seen on the Northern blot is therefore likely to correspond to these three variants. A signal could be detected in all the tissues analysed (although weak in some), suggesting ubiquitous expression of \textit{VPS13A}. This result was confirmed by RT-PCR experiments on a multiple-tissue cDNA panel and in RNA isolated from fibroblasts and a lymphoblastoid cell line (Fig. 4.1b).
**Figure 4.1 - Expression analysis of VPS13A mRNA.**

a. Adult human multiple-tissue Northern blot analysis. PCR primers KIAArt22 and KIAArt15 (see Appendix 1) were used to generate a 1028-bp cDNA probe containing *VPS13A* exons 7-17. The 2-kb β-actin probe supplied with the Northern blot (section 2.6.2) was used as control.

b. RT-PCR analysis on an adult human multiple-tissue cDNA panel and lymphoblast and fibroblast cDNA. PCR primers KIAArt13 and KIAArt11 (see Appendix 1), amplifying a region containing *VPS13A* exons 22-26, were used. The expected PCR product (563 bp) is present in all the tissues analysed.

Key: Sk, skeletal; -, PCR negative control.
4.2 Chorein-specific rabbit polyclonal antibody production

A desirable reagent for the functional characterisation of any gene product is an antibody that recognises the relevant protein. Early on in the characterisation of \textit{VPS13A}, we synthesised a peptide corresponding to amino acids 2572-2593 of chorein to raise a rabbit polyclonal antibody. However, the specificity of detection was poor (data not shown). We therefore decided to generate an antibody against a larger region of chorein, in the hope that the increased number of potential epitopes in the antigen would result in a more effective antibody.

4.2.1 Expression of GST-chorein fusion proteins

I assembled three bacterial expression vectors containing cDNA fragments encoding chorein amino acids 27-326, 1189-1502 & 2756-3055, fused to glutathione-S-transferase (GST), according to the scheme illustrated in Figure 4.2a. Each construct (pGEX-chac1, pGEX-chac2 and pGEX-chac3) was transformed into a bacterial strain engineered for efficient, inducible recombinant protein expression. Duplicate cultures of transformants were grown to log phase at 37°C, and protein expression was then induced at 18, 30 or 37°C. Expression levels of the fusion protein were investigated by SDS-PAGE at two, four and six hours post induction. As shown in Figure 4.2b, maximal expression of GST-chor1 was detected after six hours induction at 37°C; maximal expression of GST-chor2 was detected after six hours induction at either 30 or 37°C. I was not able to detect GST-chor3 expression (Figure 4.2b, bottom panel), so continued with GST-chor1 and GST-chor2 only.
Figure 4.2 - Expression of GST-fusion proteins in bacteria. a, Strategy for construction of pGEM-chac1, pGEM-chac2 and pGEM-chac3 (pGEX-chacx) and expression of GST-chorx. Three regions of VPS13A cDNA (RT1, nucleotides 79-978; RT2, nucleotides 3568-4507; RT3, nucleotides 8269-9166) were amplified using primers containing 5' XhoI sites (primers: XhoRTIF & -R, XhoRT2F & -R, XhoRT3F & -R; see Appendix 1). Purified PCR products were ligated into the TA cloning vector, pGEM-T. Inserts of the resultant constructs were fully sequenced before being liberated by XhoI digestion and ligated into the bacterial expression vector, pGEX-4T1. Ligation junction sequences were sequenced and correct constructs were transformed into BL21 bacteria for IPTG-inducible expression of fusion protein. b, Induction trials for GST-chorx expression. Crude bacterial lysates prepared 2, 4 & 6 h after induction at different temperatures were subjected to electrophoresis on a 10% polyacrylamide gel, followed by staining with Coomassie blue. Arrows indicate bands corresponding to GST-fusion protein (predicted molecular weight ~61 kDa).

Key: X, XhoI site; Ptac, IPTG-inducible promoter; Amp', ampicillin resistance gene; *, 6h uninduced control; M, size marker.

Fig 4.2
4.2.2 Fractionation of GST-chorl & -chor2 crude lysates

Overexpression of recombinant protein in bacteria can lead to incorrect folding and aggregation into insoluble inclusion bodies (Frangioni and Neel 1993). Indeed, when I induced expression of chorein fusion protein at 37°C, it was all found in the insoluble fraction (see Figure 4.3a). I therefore investigated the effects of different induction temperatures on the solubility of GST-chorl and GST-chor2 fusion proteins. I also attempted to influence protein solubility by addition of 6% ethanol upon induction. This is thought to encourage overproduction of chaperones, which can facilitate correct folding of the recombinant protein (Thomas and Baneyx 1997). However, as shown in Figure 4.3b, lowering of induction temperature did not have a measurable effect on GST-chorl solubility; and when ethanol was added little or no fusion protein was detected. Similar results were obtained with GST-chor2; and even when induced overnight at 16°C, no soluble GST-chorl or GST-chor2 was detected (data not shown). My next strategy was therefore to solubilise the fusion protein post extraction. As overall levels of GST-chorl expression were higher than GST-chor2 (Figure 4.2b), I decided to proceed with the former.

4.2.3 Solubilisation and purification of chorl antigen

To maximise expression of GST-chorl, I induced the pGEX-chacl transformant culture overnight at 37°C. Solubilisation of the fusion protein by addition of sarkosyl and purification on glutathione-agarose beads was achieved largely as described by Frangioni and Neel (1993). However, it was necessary to include detergent in all bead washing and fusion protein cleavage steps to prevent precipitation of chorl. Duplicate small-scale purifications using 0.01-0.9% sarkosyl, 2% Triton X-100, 2% Tween-20, 2% n-octyl, β-D-glucopyranoside and 1-5 mM DTT in various combinations
established that the optimal post-binding buffer was 0.1% sarkosyl, 1 mM DTT (data not shown). As shown in Figure 4.3c, the eluates (E1-E4) were not entirely pure – as well as chorl, they contained uncleaved GST-chorl and GST. The estimated purity of the pooled chorl eluate was 73%: attempts to increase this purity by, e.g., decreasing the thrombin cleavage time, or reincubating the post-cleavage eluate with glutathione-agarose beads, were unsuccessful (data not shown). Sarkosyl, a strong ionic detergent, partially denatures proteins, so its presence may have destabilised the GST-glutathione binding. Eluates E1-E3 (Figure 4.3c) were pooled and used to raise anti-chorl polyclonal antisera in rabbits SK274 and SK275.

4.3 Testing of anti-chorl antisera

I initially tested the specificity of the anti-chorl antisera 59 days post-immunisation by dot blot of chorl antigen. Comparison with preimmune sera showed that antisera from both rabbits were specific, and able to detect 50 ng of chorl even when diluted to 1:10000, as shown in Figure 4.4.
Figure 4.3 - Expression and purification of GST-chorl. Samples were separated on a 10% polyacrylamide gel, followed by staining with Coomassie blue. a, Fractionation of recombinant GST-chorl. Bacteria transformed with pGEX-chac1 were grown to log phase and induced at 37°C for 6 h before being centrifuged: the pellet was separated into soluble and insoluble fractions. b, Effect of ethanol and temperature on GST-chorl solubility. Bacterial cultures were induced for 6 h at 18 or 30°C in the presence (+) or absence (-) of 6% ethanol, and were then fractionated. c, Samples from various stages of the purification of chorl from bacteria expressing GST-chorl (for details see section 2.12.4). Bands corresponding to GST-chorl, chorl and GST are arrowed in orange, red and yellow, respectively. Key: W, whole crude bacterial lysate; S, soluble fraction; I, insoluble fraction; E1-E4, eluates 1-4; B, glutathionesepharose beads post-thrombin cleavage; BSA, bovine serum albumin standards.
Figure 4.4 - Dot blot analysis of anti-chol antiseraum.Dot blotting was performed as detailed in Materials and Methods, using pre-immune serum and serum 59 days post-chol immunisation, both from rabbit SK274. Similar results were obtained with rabbit SK275 (data not shown).
4.3.1 Detection of exogenous chorein

Figure 4.5a & b shows Western blot analysis of lysates from mammalian cells transfected with different chorein-fusion expression constructs. pcD4e encodes a 47-kDa N-terminal region of chorein. When probed with anti-chorl antiserum, the insoluble fraction of cells transfected with this construct (Figure 4.5a, lane 3) shows a single band of approximately 50 kDa that is absent in untransfected cells (lanes 1 & 2). FITR293-CAG-1.1 is a cell line stably transfected with a construct encoding full-length chorein fused at its C-terminus to enhanced green fluorescent protein (EGFP). This fusion protein has a predicted Mw of 390 kDa. In these cells, a broad band > 220 kDa was detected by anti-chorl (Figure 4.5b, lane 1). Interestingly, a slightly smaller band was seen in cells that were not expressing recombinant protein (lanes 2-5).

4.3.2 Detection of endogenous chorein

In order to explore whether the high-molecular-weight band detected in untransfected cells was indeed chorein, I prepared lysates from ChAc patient (P) and control (C) lymphoblastoid cell lines. These were analysed by Western blot using anti-chorl antiserum, as shown in Figure 4.5c. A band running above the 220-kDa marker was detected in control cells (lane 5) - this is consistent with the predicted molecular weight of chorein, which is 360 kDa. P is the proband from Wave I family CHAC2. As discussed in section 3.1.1, CHAC2 patients are homozygous for the frameshift mutation 1592del that leads to a premature stop codon in exon 17 of the VPS13A gene. They are therefore predicted to be unable to express full-length chorein. As Figure 4.5c shows, the high molecular-weight band seen in control cells was absent in P cells (lane 6). This signal was not detected using pre-immune serum (lanes 1 & 2),
or when using serum depleted in chor1-binding antibodies (lanes 3 & 4), thereby confirming the specificity of the detection.

4.4 Purification of anti-chor1 antibody

In an attempt to further increase detection specificity, I purified chor1-binding antibodies from the anti-chor1 antiserum by antigen affinity purification, as shown in Figure 4.6a. Although the purification was successful and the optimum dilution of antibody for endogenous chorein was determined, there appeared to be little difference in specificity between the antibody and the antiserum from which it was purified (Figure 4.6b). Subsequent analyses confirmed this view (data not shown), so we decided to continue using the antisera for Western blot detection.
Figure 4.5 - Detection of exogenous and endogenous chorein. 

**a,** Western blot analysis of transiently-transfected cells. COS-1 cells were either transfected with N-terminal chorein construct pcD4e (+), or were mock-transfected (-). Cell lysates were prepared 48 h post-transfection and subjected to electrophoresis on a 10% polyacrylamide gel. Immunodetection was performed with a 1:30000 dilution of rabbit SK275 serum 59 days post-chorl immunisation. 

**b,** Western blot detection of exogenous and endogenous chorein. Lysates were prepared from Flp-In T-Rex 293 cells (Invitrogen) stably transfected with full-length chorein construct (cell line FITR293-CAG-1.1, +), or from untransfected 293T cells (-). Lysates were separated on a 3-8% polyacrylamide gradient gel; immunodetection was performed with a 1:5000 dilution of rabbit SK275 serum 87 days post-chorl immunisation. Bands corresponding to chorein fusion proteins are arrowed. 

**c,** Specificity of chorein detection by western blot analysis. 20-µg protein samples from lymphoblastoid cell lines were separated by SDS-PAGE and analysed by western blot. Blots were analysed with preimmune serum (lanes 1 & 2) or anti-chorl antiserum (lanes 3-6). Anti-chorl antiserum was either run through a column bound with chorel antigen to remove chorel-binding antibodies (+, lanes 3 & 4) or left untreated (-, lanes 5 & 6). Immunodetection of early endosome antigen 1 (EEA1) shows equal loading of samples. 

Key: I, insoluble fraction; S, soluble fraction; C, healthy control; P, proband of family CHAC2.

*Construct pcD4e and stable cell line FITR293-CAG-1.1 were kindly provided by Antonio Velayos.*
Figure 4.6 - Purification of anti-chorl antiserum. a, Samples from various stages of the purification of anti-chorl antibodies from rabbit SK274 serum 87 days post-chorl immunisation. Serum was purified by affinity chromatography using a column containing bound chorl antigen (details of procedure in section 2.18.3). 16 μl of a 1:10 dilution of pre-elution column wash 1 (PrW1) and neat aliquots of pre-elution wash 2 (PrW2), eluates 1-6 (E1-6) and post-elution column wash 1 (PoW1) were subjected to electrophoresis on a 12.5% polyacrylamide gel, which was then stained with Coomassie blue. Eluates containing purified anti-chorl antibodies were subsequently pooled into two fractions, E1-3 and E4-6. b, Titration of purified E1-3 dilutions for Western blot analysis. Lysates prepared from control (C) and ChAc patient (P) lymphoblastoid cell lines were subjected to electrophoresis on a 3-8% polyacrylamide gel. Immunodetection was performed with different dilutions of E1-3, E4-6 and crude anti-chorl antiserum. The band corresponding to chorein is arrowed; a ~70-kDa band that cross-reacts strongly with anti-chorl is asterisked. Optimal dilution of E1-3 appears to be 1:1000. Similar results were obtained for purification of anti-chorl antibodies from rabbit SK275 antiserum, except that the optimal dilution for E1-3 was 1:2500.
4.5 Analysis of chorein distribution

Lysates were prepared from a variety of cell lines routinely used in tissue culture techniques. Human cell lines were derived from cervical carcinoma (HeLa), foetal lung (MRC-5 SV2), embryonic kidney (293T), lymphoblasts (lymph), neuroglioma (H4), rhabdomyosarcoma (RD), hepatocarcinoma (Hep3B) and myelogenous leukaemia (K562). Lysates were also prepared from two non-human cell lines (COS-7, from African green monkey kidney; and CHO-K1, from Chinese hamster ovary). All 10 cell-line lysates were analysed by Western blot. As shown in Figure 4.7, antichor1 antiserum detects a high molecular weight signal in all cell lines analysed. A much fainter signal was detected in the non-human cell lines (lanes 1 & 2). This implies that these cells express much less chorein than the others do, or that there is less cross-reaction of the non-human chorein homologues with the anti-chor1 antiserum.
Figure 4.7 - Expression of chorein in different cell types. 20-μg protein samples from different cell lines were separated by SDS-PAGE and analysed by western blot. Cells were derived from African green monkey (lane 1), Chinese hamster (lane 2), or human (lanes 3-10) tissue. Immunodetection of early endosome antigen 1 (EEA1) shows equal loading of samples.
4.6 Expression of mutant chorein

4.6.1 Expression of mutant chorein in ChAc cell lines

Figure 4.8a shows a Western blot of lymphoblastoid protein lysates prepared from individuals from two ChAc pedigrees. As described in section 3.1.1, the affected siblings in the CHAC6 pedigree are homozygous for the missense mutation 4354T>C in exon 37 of VPS13A, which results in substitution of a proline for a serine residue at amino-acid position 1452 in chorein. Their unaffected brother is homozygous for the wild type. The affected son in family CHAC11 has inherited the 237del mutation in exon 4 from his mother and the 9429_9432del mutation in exon 72 from his father. In contrast to the unaffected individuals in both pedigrees, CHAC6 and CHAC11 patients (lanes 2, 4 & 7) show a markedly reduced level of chorein expression compared with the control. The complete absence of a corresponding band in the CHAC2 patient (P, lane 1) implies that the weak staining seen in the other patients is indeed chorein and not an unspecific signal. The 237delT mutation in family CHAC11 is predicted to lead to degradation of most of the mutant transcript by nonsense-mediated decay. Any protein that is produced from this allele will be severely truncated (predicted size 10 kDa). The faint signal seen for the proband must therefore derive from truncated protein generated by the 9429_9432del mutation (357 kDa); and/or chorein isoform 1B (345 kDa), which is encoded by exons 1-69 only and is theoretically unaffected by the exon 72 mutation.

As MLS can be difficult to distinguish clinically from ChAc (Danek et al. 2004), lymphoblasts from an MLS patient with the XK splice-site mutation 508+1G>A (FD in (Ho et al. 1994)) were also analysed to determine whether chorein detection could
discriminate between the two disorders. As shown in Figure 4.8a (M1, lane 8), the level of chorein expression in this patient was similar to that of the control.

Chorein expression in ChAc patient and control primary skin fibroblasts is shown in Figure 4.8b. Wave II patient 24 (II-24) is a compound heterozygote for the nonsense mutation S2140X in exon 48 and the frameshift mutation 9190del in exon 70.

Expression of chorein in this patient is detectable (lane 1), but markedly reduced when compared with fibroblasts from a healthy control (lane 4). II-11 is heterozygous for the frameshift mutation 1125_1128del in exon 13; this was the only mutation detected upon screening of this patient's VPS13A gene (Table 3.2). II-2 is a compound heterozygote for two splice-site mutations: 2288+2T>C in intron 22 and 8472-1G>C in intron 61. Expression of chorein in these latter two patients was undetectable (lanes 2 & 3).
Figure 4.8 - Analysis of chorein expression in ChAc cell lines. **a**, Analysis of lymphoblastoid cell lines. 20-μg protein samples from lymphoblastoid cell lines were separated by SDS-PAGE and analysed by western blot. Cells were derived from family CHAC2 proband (P, lane 1), affected (lanes 2, 4 & 7) and unaffected (lanes 3, 5 & 6) members of CHAC6 and CHAC11 families, an individual with McLeod syndrome (M1, lane 8) and a healthy control (C, lane 9). The VPS13A genotype is indicated above CHAC6 and CHAC11 family members (+, wild type; -, mutant allele). Immunodetection of early endosome antigen 1 (EEA1) shows equal loading of samples. **b**, Analysis of fibroblasts. 20-μg protein samples from primary skin fibroblasts derived from ChAc patients (lanes 1-3) and a control (C, lane 4) were separated by SDS-PAGE and analysed by western blot. Immunodetection of early endosome antigen 1 (EEA1) shows equal loading of samples. The panel on the right shows a longer exposure of the anti-chorein immunoblot.
4.6.2 Expression of chorein in erythrocyte membranes

Erythrocyte membrane fractions were prepared from frozen whole blood from seven unrelated ChAc patients and two healthy individuals (C1 & 2). Blood from an MLS patient (M2), who harbours a hemizygous deletion of XK exon 1 (Patient 13 in (Danek et al. 2001)), was also used. Blood had been stored at -20°C for different lengths of time, ranging from 8 weeks 4 days (C2) to 159 weeks 5 days (II-23). M2 blood had been stored at -192 °C for 654 weeks 3 days. Western blot analyses of these preparations are shown in Figure 4.9.

The ChAc patients were known to harbour at least one heterozygous mutation likely to disrupt VPS13A (Tables 3.2 & 3.3). As shown in Figure 4.9, chorein expression was not detected in erythrocyte membrane preparations from patients III-29, III-33, III-26 and II-23 (lanes 2, 3, 5 & 7). Expression of chorein in patient III-8 was detectable (lane 6), but markedly reduced when compared with controls C3 and C4. Overexposure of the immunoblot revealed a very faint signal corresponding to chorein for patients II-16 & III-7 (data not shown). For patients II-16, III-29 and III-33 the reduction or absence of signal cannot be explained by degradation of the chorein protein during storage, as they were stored for less than, or the same amount of time as, C1. Although their blood was stored for longer, this is also an unlikely explanation for reduced signal in the remaining patients, as evidenced by the integrity of the cross-reacting signal at ~220 kDa. A signal comparable in strength with the controls was observed in the MLS patient (lane 8), even though this blood had been stored for over 12 years.
Figure 4.9 - Analysis of chorein expression in red blood cells. Red blood cell membranes were prepared from ChAc patients (lanes 1-7), an individual with McLeod syndrome (M2, lane 8) and healthy controls (C1 & 2, lanes 9 & 10). Membrane preparations were separated by SDS-PAGE and analysed by western blot. The band corresponding to chorein is arrowed - the lower band (~220 kDa, asterisked) probably results from cross-reaction of anti-chor1 with spectrin, the most abundant protein associated with the erythrocyte membrane (Marchesi et al., 1976). Total protein staining of the blot prior to immunodetection demonstrated equal loading of samples (data not shown).
4.7 Summary

This chapter describes the analysis of VPS13A gene expression, at both the mRNA and protein levels. mRNA expression was analysed by Northern blot and RT-PCR, and was shown to be ubiquitous. For analysis of protein expression, an N-terminal region of chorein (chorl) was expressed in bacteria and purified. This was used to raise rabbit polyclonal antisera (anti-chorl). Western blot analyses using anti-chorl suggested that chorein expression was also ubiquitous. However, chorein expression was absent or markedly reduced in lymphoblastoid cell lines, primary fibroblasts and erythrocyte membranes prepared from ChAc patients. In contrast, lymphoblastoid cell lysates and erythrocyte membrane preparations from MLS patients showed chorein levels similar to those of healthy controls.
Chapter 5: Detection of potential chorein-interacting proteins

5.1 Yeast two-hybrid screen of chorein

The yeast two-hybrid (Y2H) system is an *in vivo* technique whereby interactions between two proteins are identified through reconstitution of an active transcription factor (TF) (reviewed in (Toby and Golemis 2001)). Yeast are transformed with two fusion proteins: a TF DNA-binding domain fused to a protein of interest, A (the 'bait'), and a TF activation domain fused to a second protein, B (the 'prey'). If proteins A and B interact, the two domains of the transcription factor are brought together and transcription of reporter genes containing the relevant TF-binding site is activated. This technique can be used to reveal interactions between two known proteins, or to screen proteins encoded by a cDNA library for those that interact with a protein of interest. I used a GAL4 Y2H system from Invitrogen to search for proteins that interact directly with chorein.

5.1.1 Construction of yeast two-hybrid baits

I assembled six Y2H bait constructs containing cDNA fragments encoding chorein amino acids 1-600, 1961-2558, 2550-3174, 495-1096, 993-1597 or 1487-2090, fused to the DNA-binding domain of GAL4 (GAL4DB), according to the scheme illustrated in Figure 5.1a & b. The resultant constructs (pDEST32chacY1-6, respectively) were transformed into yeast and cultures were made of the transformants. Western blot analysis of these cultures revealed that each was expressing a GAL4DB-fusion protein of the appropriate size, although expression levels were variable (Figure 5.1c).
5.1.2 *Investigation of potential homotypic interactions*

I investigated the possibility that chorein fragments could form interactions with themselves. I constructed six Y2H prey constructs (pDEST22chacY1-6) encoding the same chorein regions as used for the baits, but this time fused to GAL4AD (see Figure 5.2a). Each bait and prey construct pair was transformed into yeast strain MaV203, which contains *HIS3*, *URA3*, and *lacZ* reporter genes. When activated, *HIS3* allows growth on media lacking histidine (-His), *lacZ* results in a blue colour when assayed with X-Gal, and *URA3* permits growth on media lacking uracil (-Ura) but inhibits growth on media containing 5-fluoro-orotic acid (+5FOA). Double transformants were assayed for *lacZ* activity and for growth on the relevant selection media. As shown in Figure 5.2b, no *lacZ* activity was seen, and colonies grew on +5FOA but not on -His or -Ura: this pattern of growth matched that of the negative interaction control.
Figure 5.1 - Construction and expression of baits for yeast two-hybrid analysis. 

**a**, Schematic representation of chorein primary structure. Regions of high conservation between chorein homologues are indicated in blue; vertical bars indicate putative transmembrane domains; ellipses indicate putative tetratricopeptide repeats (discussed in section 1.8). Horizontal bars below chorein indicate the PCR products amplified from VPS13A cDNA for bait construction; the amino acids encoded by each bait are indicated in parentheses. 

**b**, Strategy for construction of pDEST32chacY1-6. The regions of interest were amplified using PCR primers with 5' attB recombination site sequences (see Appendix 1). Recombination reactions with the PCR products and pDONR vectors generated entry clones pENTRchacY1-6. The gene regions in the entry clones were transferred into the destination vector pDEST32, generating bait vectors pDEST32chacY1-6. Regions B1&2, P1&2, L1&2 and R1&2 are recombination sites. 

**c**, Western blot analysis of bait fusion proteins. 30 µl of crude cell lysate from yeast transformed with pDEST32chacY1-6 (see section 2.13.1) were subjected to electrophoresis on a 10% polyacrylamide gel. Immunodetection was performed with anti-GAL4 DB antibody.

**Key**: ccdB, inhibitor of DNA gyrase; DB, GAL4 DNA-binding domain; Km', kanamycin resistance gene; Gm', gentamicin resistance gene; LEU2, gene enabling growth on leucine-deficient media.
Figure 5.2 - Construction of prey vectors and investigation of homotypic interactions in chorein. 

a, Strategy for construction of pDEST22chacY1-6. The gene regions in the entry clones pENTRchacY1-6 were transferred into the destination vector pDEST22, generating prey vectors pDEST22chacY1-6. Bait-prey vector pairs were cotransformed into yeast strain MaV203 and double transformants were analysed for induction of reporter gene expression.

b, Analysis of homotypic interactions. Double transformants were patched on to SC-Leu-Trp media according to the scheme in the top row (master) and incubated for 16 h at 30°C. The patches were then replica plated on to selection media and incubated at 30°C. lacZ activity was assayed 24 h after replica plating (second row; growth on -His, -Ura, or +5FOA media was recorded 48 h after replica plating (third, fourth & fifth rows).

Key: ccdB, inhibitor of DNA gyrase; AD, GAL4 transactivation domain; DB, GAL4 DNA-binding domain; Km', kanamycin resistance gene; Gm', gentamicin resistance gene; Ap', ampicillin resistance gene; a, negative interaction control; b, 'weak interaction' control; c, 'moderate interaction' control; B1-P1, bait construct pDEST32chacY1 cotransformed with prey construct pDEST22chacY1; B1-P0, pDEST32chacY1 cotransformed with empty prey vector pEXP-AD502.
5.1.3 Yeast two-hybrid library screen

Y2H bait constructs pDEST32chacY1-6 were used to screen an adult human brain cDNA library cloned into the prey vector pEXP-AD502. In each case, over a million clones were screened for chorein-interacting proteins (Table 5.1). Clones that grew on media lacking leucine, tryptophan and histidine (primary positive clones) were patched on to secondary selection media, testing in parallel for activation of all three reporter genes. Clones whose growth patterns implied the presence of a chorein-interacting protein (secondary positive clones) were cultured, and prey cDNA constructs were isolated. These were retested by cotransformation with the relevant bait construct or an empty pDEST32 plasmid into a new MaV203 host. Clones that gave positive results in the new host (tertiary positive clones) were deemed worthy of further investigation. As shown in Table 5.1, only seven tertiary positive clones were identified, all from the screen using pDEST32chacY2 as the bait. The prey constructs isolated from these clones were named pEXP-c1 through to pEXP-c7. (When DNA was isolated from positive clone 4, restriction analysis revealed the presence of two populations of prey constructs - these were called pEXP-c4 and pEXP-c4a.) None of these constructs promoted activation of reporter genes in the absence of chacY2, as shown in Figure 5.3. When cotransformed with pDEST32chacY2, pEXP-c1, -c3, -c5 & -c7 showed weak lacZ activation; pEXP-c2, -c4, & -c6 showed very weak activation of lacZ. pEXP-c4a did not show detectable lacZ activation. Consistent with these results, pEXP-c1, -c3, and -c5 transformants grew well on -His media but poorly on +5FOA media. pEXP-c7 also grew poorly on +5FOA. pEXP-c2, -c4 & c6 transformants grew on -His media; inhibition of growth on +FOA was slight but detectable. pEXP-c4a did not grow on -His media and grew well on +FOA media.

Growth on -Ura media, which requires a moderately strong interaction between bait...
and prey, was not observed for any test transformant. From these results, it was concluded that pEXP-c1, -c3, -c5 & -c7 constructs encoded probable weak chacY2-interactors, while pEXP-c2, -c4 & -c6 constructs encoded proteins that probably interact very weakly with chacY2. pEXP-c4a was concluded to encode a protein that does not interact with chacY2, and was therefore excluded from further analysis.

Table 5.1 - Number of clones processed in chorein yeast two-hybrid screens

<table>
<thead>
<tr>
<th>Bait construct</th>
<th>No. clones screened</th>
<th>No. primary positives</th>
<th>No. secondary positives</th>
<th>No. tertiary positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDEST32chacY1</td>
<td>$4.8 \times 10^6$</td>
<td>203$^*$</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>pDEST32chacY2</td>
<td>$4.4 \times 10^6$</td>
<td>62</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>pDEST32chacY3</td>
<td>$4.7 \times 10^6$</td>
<td>102</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>pDEST32chacY4</td>
<td>$3.5 \times 10^6$</td>
<td>51</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>pDEST32chacY5</td>
<td>$1.2 \times 10^6$</td>
<td>35</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>pDEST32chacY6</td>
<td>$4.2 \times 10^6$</td>
<td>30</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

$^*$ The relatively large number of primary positives in this first screen probably doesn’t reflect the true total, but rather the increased ability to discard obvious false positives in subsequent experiments.
Figure 5.3 - Retransformation assays of secondary positives from chacY2 yeast two-hybrid screen.
Yeast strain MaV203 was cotransformed with bait construct pDEST32chacY2 and putative prey constructs (pEXP-c1 to pEXP-c7). Cotransformation pairs with empty bait construct pDBLeu were used as controls. Double transformants were patched on to SC-Leu-Trp media according to the scheme in the top row (master) and incubated for 16 h at 30°C. The patches were then replica plated on to selection media and incubated at 30°C. lacZ activity was assayed 24 h after replica plating (second row); growth on -His, -Ura, or +5FOA media was recorded 48 h after replica plating (third, fourth & fifth rows). Note that pDEST32chacY2/pEXP-c7 cotransformant grew poorly on -H media (third column, third row), but this result was consistent with the poor growth of interaction control B on this plate.
Key: a, negative interaction control; b, ‘weak interaction’ control; c, ‘moderate interaction’ control; B0-Pc1, empty bait vector pDBLeu cotransformed with prey construct pEXP-c1; B2-Pc1, bait construct pDEST32chacY1 cotransformed with pEXP-c1.
5.1.4 Characterisation of tertiary positives

The seven Y2H prey constructs that appeared to encode chacY2-interacting proteins were sequenced using pDEST22-specific primers. They were identified by BLAST comparison of the nucleotide sequence in the GenBank database (http://www.ncbi.nlm.nih.gov/BLAST/). I determined that pEXP-c1, -c5 & -c7 encoded part of mitogen-activated protein kinase kinase kinase 7 interacting-protein 2, also known as TAK1-binding protein 2 (TAB2). pEXP-c3, -c4 & -c6 encoded part of the putative protein phosphatase 2C eta (PP2Cη). pEXP-c1 encoded part of the p65 subunit of nuclear transcription factor kappa B (NF-κB). Sequencing revealed that the cDNA for each protein had been cloned in-frame with the GAL4 sequence.

5.1.5 Further Y2H investigation of chacY2 - PP2Cη interaction

Sequencing of pEXP-c3, -c4 & -c6 revealed that all three constructs contained sequence that was largely identical to the predicted ORF of cDNA FLJ39362,² apart from a 97-bp deletion and a 64-bp insertion. FLJ39362 bases 723-819 were deleted in the constructs; the insertion occurred between FLJ39362 bases 939 and 940, and corresponded exactly to bases 836-899 of cDNA FLJ32332. The amino-acid sequences predicted for FLJ32332 and FLJ39362 correspond exactly to PP2Cη splicing isoforms 1 and 2, respectively. Two other splicing variants are annotated for this protein - it seems likely that the inserts of pEXP-c3, -c4 & -c6 encode a fifth splicing variant of PP2Cη. In addition, pEXP-c3 contained 350 bases and pEXP-c4 & -c6 contained 320 bases upstream of the predicted initiating codon. This meant that the GAL4AD-fusion protein being expressed included either 117 aa (pEXP-c3) or 107 aa (pEXP-c4 & -c6) that were not predicted to be present in the wild-type protein. The

² GenBank accession numbers are reported in Appendix 2.
initial assay results (Fig. 5.3) appeared to show that pEXP-c3 led to stronger activation of lacZ transcription than pEXP-c4 & -c6 did. This prompted us to investigate whether the extra N-terminal amino acids predicted to be absent in the wild-type protein were contributing to the observed interaction with chacY2. I assembled a prey construct encoding GAL4AD fused to the first 117 aa only of the clone 3 protein (Figure 5.4a). This construct (pEXP-FLJ5'), along with intact pEXP-c3 & -c4, was cotransformed with pDEST32chacY2, pDEST32chacY1 or pPC97-dDP. Double transformants were assayed for HIS3 and lacZ activity. As shown in Figure 5.4b, only intact pEXP-c3 & -c4 showed activation of reporter genes in combination with chacY2 bait. They did not activate transcription when cotransformed with the spurious bait constructs pDEST32chacY2 and pPC7-dDP. This time, pEXP-c3 and pEXP-c4 transformants showed roughly equivalent levels of lacZ activation and growth on -His media.
Figure 5.4 - Investigation of the interaction between chacY2 and PP2C\eta. a, Strategy for construction of pEXP-FLJ5'. Prey construct pEXP-c3 was digested with MscI and SacI to remove the coding sequence of putative protein PP2C\eta (the predicted initiation codon is highlighted in green). T4 DNA polymerase (T4 pol) was used to remove the 3' overhang generated by SacI digestion. The fragment was then religated to generate pEXP-FLJ5'. This construct contains only sequence 5' to the putative start site, in frame with a TAA stop codon (indicated in red). b, Analysis of interactions between PP2C\eta and chacY constructs. Yeast strain MaV203 was cotransformed with prey comprising different regions of PP2C\eta cDNA (pEXP-c3, pEXP-FLJ5' or pEXP-c4), and either 'correct' bait pDEST32chacY2 or spurious baits (pDEST32chacY1 or pPC97-dDP). Transformant pairs with empty prey vector were used as controls. Double transformants were patched on to SC-Leu-Trp media according to the scheme in the left-hand panel (master) and incubated for 16 h at 30°C. The patches were then replica plated on to selection media and incubated at 30°C. lacZ activity was assayed 24 h after replica plating (middle panel); growth on -His media was recorded 48 h after replica plating (right-hand panel).

Key: Ap', Ampicillin resistance gene; AD, GAL4 transactivation domain; a, negative interaction control; b, 'weak interaction' control; c, 'moderate interaction' control; B1-Pc3, bait construct pDEST32chacY1 cotransformed with prey construct pEXP-c3; Bdp-Pf5', bait construct pPC97-dDP cotransformed with prey construct pEXP-FLJ5'; B1-P0, pDEST32chacY1 cotransformed with empty prey vector pEXP-AD502.
5.1.6 Further Y2H investigation of chacY2 - NF-κB interaction

Sequence analysis of pEXP-c2 revealed that it contained bases 170-1767 of NF-κB cDNA, encoding amino acids 91-551 of the 551-aa p65 subunit. pEXP-c2 had not shown any activation of reporter genes when transformed with an empty bait vector (Figure 5.3). However, as NF-κB is a transcription factor itself (and the Y2H assay is based on reconstitution of an active TF), we were suspicious that its interaction here was not specific to chacY2. To investigate this possibility, I cotransformed pEXP-c2 with the specific bait construct pDEST32-chacY2 or with the spurious bait constructs pDEST32-chacY1 or pPC97-dDP. As shown in Figure 5.5, only yeast cotransformed with pEXP-c2 and the specific bait displayed significant \( \text{lacZ} \) activity and growth on -His media.

5.2 Investigation of chorein interaction with TAB2

Sequence analysis of prey constructs containing TAB2 sequence revealed that pEXP-c1 and pEXP-c5 contained nucleotides 1125-4359 of TAB2 cDNA, encoding amino acids 318-693; while pEXP-c7 contained a slightly smaller fragment, nucleotides 1182-4355 (amino acids 337-693). This represents the 357 aa at the C-terminus of TAB2. Although two splicing variants have been reported for this cDNA, all clones were derived from splice variant 2 only. I performed co-immunoprecipitation and colocalisation studies in order to further characterise the interaction between TAB2 and chorein.
Figure 5.5 - Investigation of the interaction between chacY2 and NF-κB. Yeast strain MaV203 was cotransformed with NF-κB construct pEXP-c2 and either 'correct' bait pDEST32chacY2 or spurious baits (pDEST32chacY1 or pPC97-dDP). Transformant pairs with either empty bait or empty prey vector were used as controls. Double transformants were patched on to SC-Leu-Trp media according to the scheme in the top panel (master) and incubated for 16 h at 30°C. The patches were then replica plated on to selection media and incubated at 30°C. lacZ activity was assayed 24 h after replica plating (middle panel); growth on -His media was recorded 48 h after replica plating (bottom panel).

Key: a, negative interaction control; b, 'weak interaction' control; c, 'moderate interaction' control; B1-P0, bait construct pDEST32chacY1 cotransformed with empty prey vector pEXP-AD502; Bdp-P0, bait construct pPC97-dDP cotransformed with pEXP-AD502, B0-Pc2, empty prey vector pDBLeu cotransformed with prey construct pEXP-c2.
5.2.1 Co-immunoprecipitation of chorein and TAB2

I assembled the mammalian expression construct pcD4-chac2 according to the strategy set out in Figure 5.6a. pcD4-chac2 contains the cDNA encoding amino acids 1959-2558 of chorein fused to a C-terminal myc epitope, under control of a strong human cytomegalovirus immediate-early (CMV) promoter. pCMVT-TAB2, a kind gift from Kunihiro Matsumoto (Nagoya University, Japan), encodes full-length TAB2 fused to an N-terminal T7 epitope, also under control of the CMV promoter. 293T cells were cotransfected with pcD4-chac2 and pCMVT-TAB2. Cell extracts were immunoprecipitated with a monoclonal antibody to T7, and coprecipitated chac2-myc was detected by immunoblotting with an anti-myc antibody (Figure 5.6b, panels i-iii). These results were suggestive of an association between TAB2 and chac2. However, when the experiment was reversed and immunoprecipitation was performed with anti-myc antibody, coprecipitated T7-TAB2 could be detected even in the absence of chac2-myc (panel iv, lane 3), implying an unspecific interaction directly between T7-TAB2 and the anti-myc antibody. Repeating this experiment gave the same results. In an attempt to gain more unambiguous data for this interaction, I repeated the coimmunoprecipitation, this time using full-length chorein.
Figure 5.6 - Co-immunoprecipitation of T7-TAB2 and chac2-myc. a, Strategy for construction of pcD4-chac2. The region encoding amino-acids 1959-2558 of chorein was amplified from VPS13A cDNA using a forward primer containing a 5' KpnI and a reverse primer containing a 5' XhoI site (Kpnchac2F & Xhochac2R: see Appendix 1). Purified PCR products were ligated into the TA cloning vector, pGEM-T. Inserts of the resultant constructs were fully sequenced before being liberated by KpnI/XhoI digestion and ligated into the mammalian expression vector, pcDNA4-TO-mycHA. Ligation junction sequences were sequenced and correct constructs were used in cotransfection studies with pCMVT-TAB2 (schematic shown on the right). b, Coimmunoprecipitation of TAB2 and chac2. HEK293T cells were transfected with constructs according to the following schedule: 1 - no constructs; 2 - pcD4-chac2; 3 - pCMVT-TAB2; 4 - pcD4-chac2 & pCMVT-TAB2. Cells were lysed 48h post-transfection and the lysates immunoprecipitated (IP) with anti-myc or anti-T7 antibody. Cell lysate (no IP) and IP proteins were detected by immunoblotting (IB) with anti-myc or anti-T7. Bands corresponding to chac2-myc or T7-TAB2 are asterisked. Construct pCMVT-TAB2 was kindly provided by K Matsumoto, Nagoya University, Japan.
pcDFchacS is a mammalian expression vector that encodes full-length chorein fused to a C-terminal FLAG epitope, under control of the CMV promoter (see Figure 5.7a). I cotransfected 293T cells with pcDFchacS and pCMVT-TAB2. Cell extracts were immunoprecipitated with either an anti-T7 antibody or an anti-FLAG antibody. Coprecipitated chorein-FLAG and T7-TAB2 were detected by immunoblotting with anti-T7 antibody and anti-FLAG antibody, respectively (Figure 5.7b). Chorein-FLAG was found to coprecipitate using anti-T7 antibody only in the presence of T7-TAB2 (panel ii, lane 4). In addition, T7-TAB2 was found to coprecipitate using anti-FLAG antibody only in the presence of chorein-FLAG (panel iv, lane 4).

5.2.2 Immunolocalisation of chorein-FLAG and T7-TAB2

In order to determine the biological significance of the chorein-TAB2 interaction, I investigated the subcellular localisation of exogenous chorein and TAB2. 293T and MRC-5 SV2 cells were cotransfected with chorein-FLAG and T7-TAB2. Immunofluorescence staining revealed that the transfected proteins did not appear to colocalise, as shown in Figure 5.8a. Cells transfected only with T7-TAB2 were also investigated. Immunolocalisation of endogenous chorein using anti-chorl antiserum gave a high background signal. However, a minority of MRC-5 SV2 cells showed vesicular staining similar to that seen for transfected chorein (Figure 5.8b, panel v): this did not appear to colocalise with transfected T7-TAB2 (panel vi). Unfortunately, in the absence of an anti-TAB2 antibody, it was not possible to determine the subcellular localisation of endogenous TAB2.
Figure 5.7 - Co-immunoprecipitation of T7-TAB2 and chorein-FLAG. a, Schematic of the mammalian expression vector pcFchac8, encoding full-length chorein fused to a C-terminal FLAG epitope, under control of a CMV promoter. For schematic of pCMVT-TAB2, see Figure 5.6a. b, Coimmunoprecipitation of TAB2 and chorein. HEK293T cells were transfected with constructs according to the following schedule: 1 - no constructs; 2 - pcFchac8; 3 - pCMVT-TAB2; 4 - pcFchac8 & pCMVT-TAB2. Cells were lysed 48 h post-transfection and the lysates were immunoprecipitated (IP) with anti-FLAG or anti-T7 antibody. Cell lysate (no IP) and IP proteins were detected by immunoblotting (IB) with anti-FLAG or anti-T7. Bands corresponding to chorein-FLAG or T7-TAB2 are asterisked. Construct pcFchac8 was kindly provided by A Velayos.
Figure 5.8 - Subcellular localisation of TAB2 and chorein. a, Localisation of T7-TAB2 and chorein-FLAG. MRC5.V2 (i & iii) and HEK293T (ii & iv) cells were grown on a coverslip and cotransfected with pcFchac8 and pCMVT-TAB2. Forty hours post-transfection, cells were fixed by incubation in 4% paraformaldehyde (PFA, panels i & ii) or methanol (iii & iv). b, Localisation of T7-TAB2 and endogenous chorein. MRC5.V2 cells were mock transfected (v) or transfected with pCMVT-TAB2 (vi). Forty hours post-transfection, cells were methanol-fixed.

Rabbit anti-FLAG (a) or rabbit anti-chorl (b) and mouse anti-T7 (a & b) antibodies were used, followed by secondary staining with Alexa Fluor 594-conjugated goat anti-rabbit (red) and Alexa Fluor 488-conjugated goat anti-mouse (green). Cells were viewed by confocal fluorescent microscopy. Scale bar, 20 μm.

Fig 5.8
5.3 Summary

This chapter describes the search for proteins that may interact with chorein. This was attempted via a yeast two-hybrid screen of a human brain cDNA library, using six different regions of chorein (chacY1 - chacY6) as bait. After elimination of false positives, only seven independent prey constructs (pEXP-c1 - pEXP-c7) remained, all showing a putative interaction with chacY2. pEXP-c2 encoded a region of the p65 subunit of NF-κB: although this is itself a transcription factor, retransformation of this construct with spurious baits implied that this positive was unlikely to be an artefact of the assay. pEXP-c3, -c4 & -c6 encoded part of a putative protein phosphatase and contained some sequence predicted to be upstream of the initiating codon. Retransformation assays with a construct containing only this upstream sequence revealed that this region was not responsible for the chacY2 interaction. pEXP-c1, -c5 & -c7 encoded a C-terminal region of TAB2, a protein that interacts with mitogen-activated protein kinase kinase kinase 7. Coimmunoprecipitation studies in mammalian cells supported an interaction between TAB2 and chorein. However, immunofluorescent microscopy of cells expressing epitope-tagged chorein and TAB2 revealed that these exogenous proteins did not appear to colocalise.
6.1 Mutational analysis of *VPS13A*

6.1.1 *VPS13A* is the gene mutated in chorea-acanthocytosis

This laboratory previously located the *VPS13A* locus to a 6-cM region on chromosome 9q21-q22, in a genomewide scan of 11 families with ChAc (Rubio *et al.* 1997). After identification of a novel gene, *VPS13A*, in the ChAc critical region, we initially searched for *VPS13A* mutations in these 11 families (Wave I, (Rampoldi *et al.* 2001)). We found 16 different mutations, identifying at least one mutation in each family. Every mutation co-segregated with the disease haplotype, as confirmed by sequencing or restriction analysis of family members. Of these mutations, 12 lead to a PTC, and are therefore likely to severely disrupt gene function: this demonstrates that *VPS13A* is the gene that, when mutated, causes ChAc.

We followed up this initial analysis with two additional mutation screens of 43 (Wave II, (Dobson-Stone *et al.* 2002)) and 29 (Wave III) unrelated ChAc probands. In total, we have identified 88 different mutations likely to cause chorea-acanthocytosis in 72 pedigrees. Seventy-six of these disease mutations were only identified in a single family each, indicating a strong allelic heterogeneity with no single mutation causing the majority of ChAc cases in the population. For 11 probands in this study with typical symptoms, no disease mutations in *VPS13A* were detected. As mentioned previously (section 1.5.2), a pedigree with autosomal dominant inheritance of a

---

3 Originally reported as *CHAC* (Rampoldi *et al.*, 2001)
ChAc-like phenotype was found to be associated with the Huntington's disease-like-2 expansion in the \textit{junctophilin}-3 gene (Walker \textit{et al.} 2003). We therefore screened for this expansion in all ChAc probands in this study without identified \textit{VPS13A} mutations: however, all allele sizes observed were within the wild-type range (data not shown).

As there is little evidence for additional locus heterogeneity of ChAc, it could be that the 'missing' mutations are located within the 5' or 3' UTRs, or within intronic sequences not screened in this study; deletions and rearrangements preventing amplification of one allele will also not be detected with our screening method. For the 12 probands with only one heterozygous mutation detected, this is almost certain to be the case. Since seven deletions of entire exons have already been identified ((Ueno \textit{et al.} 2001) and work described in this thesis) and the \textit{VPS13A} gene is spread over such a large genomic region (about 240 kb), it is conceivable that many ChAc patients could harbour heterozygous deletions. Such deletions will only be detected by comprehensive Southern blot analysis of the entire gene or quantitative PCR.

6.1.2 Mode of inheritance of ChAc

Although the majority of reported ChAc pedigrees are consistent with autosomal recessive inheritance, a few studies have reported autosomal dominant (AD) inheritance of ChAc (Levine \textit{et al.} 1968; Marson \textit{et al.} 2003; Saiki \textit{et al.} 2003). In this study, two ChAc pedigrees with apparent AD inheritance were analysed (II-5 & II-8). In family II-8, we found that the father, who displayed some mild dyskinesia, was homozygous for the splicing mutation (9474G>A) he passed on to his affected daughter (Figure 3.2b). Perhaps this mutation permits some correct splicing of
transcript and it is only in combination with an additional mutation (observed in
proband II-8) that the full-blown ChAc phenotype is seen. In family II-5, we found
that the proband and his affected mother and sister were all homozygous for the same
nonsense mutation (Figure 3.2a). The mother and father are first cousins - the
mutation results, coupled with haplotype analysis showing homozygosity of the
CHAC critical region in the mother (data not shown) imply that the maternal
grandparents were also related. Therefore, both families in fact display
pseudodominant inheritance. Saiki et al. (2003) found that the proband and her
affected sibling in an AD ChAc pedigree were heterozygous for a splicing mutation in
VPS13A exon 57, which was absent in the unaffected mother. They were not able to
screen the presumptively affected father. It is therefore conceivable that, as in II-8, the
father was homozygous for this mutation, an additional undetected mutation being
inherited from the heterozygous mother. By extension, pseudodominant inheritance is
a possible explanation for other cases of apparent AD transmission: this highlights the
need for careful genetic analysis of all such pedigrees.
6.1.3 VPS13A mutation distribution

Figure 6.1 shows that there does not seem to be any ‘clustering’ of mutations identified in \textit{VPS13A} so far. However, conclusions can be drawn about the relative importance of its various splice forms. As mentioned previously (section 1.7.3), the two main splice forms as identified by expressed sequence tag analysis are isoform 1A, the result of splicing exons 1-68 and 70-73; and isoform 1B, from exons 1-69 only. Ten mutations in this study (Tables 3.1-3.3) are predicted to affect only transcript 1A. Four probands are in fact homozygous for transcript 1A-specific mutations: probands II-42, III-6 & III-24 (deletion of exons 70-73) and proband III-18 (a 2-bp deletion in exon 72). Given that these patients display symptoms typical for ChAc, we can deduce that isoform 1B, and other splice forms lacking exons 70-73, cannot compensate for absence of full-length chorein in this disorder.
**Figure 6.1 - Spectrum of VPS13A mutations identified in this study.** The coding sequence of VPS13A transcript 1A is illustrated in the middle, with exon junctions indicated by vertical lines. The number of probands in which identical mutations were identified is shown in parentheses.
6.1.4 Recurrent VPS13A mutations

Twelve mutations reported here were found in multiple probands within this study, as shown in Table 6.1. Probands II-12 and II-39 are heterozygous and II-27 is homozygous for the mutation 622C>T (R208X), which was also found in Wave I family CHAC3. Family CHAC3 and proband II-12 both originate from Italy and so it is conceivable that their mutation was inherited from a common ancestor. However, II-27 originates from Japan and II-39 has inherited the mutation from a Danish mother. As it is unlikely that these individuals are related to the other pedigrees, it seems logical to conclude that the 622C>T mutation has arisen at least three times independently. Supporting this hypothesis is the fact that the cytosine involved is part of a CG dinucleotide and as such is especially vulnerable to methylation and spontaneous deamination to thymine (Cooper and Youssoufian 1988). The same explanation is appropriate for the probable recurrence of the 9109C>T (R3037X) mutation in the Mexican family CHAC7 and the German proband II-1, and the 4411C>T (R1471X) mutation in Japanese II-26 and Italian III-3. This mechanism may also explain the overrepresentation of R>X substitutions in this study (9/22 nonsense mutations). Proband II-34 and CHAC11 both originate from the UK and therefore inheritance of the 9429_9432del allele from a common ancestor cannot be ruled out. The deletion is of the tetranucleotide AGAG within the tandem repeat AGAGAG, however, and so it is plausible that this mutation has arisen twice through replication slippage. The existence of a smaller deletion within the same region in II-29 & III-18 (9431_9432del, deletion of the dinucleotide AG) supports the latter hypothesis. The duplication of the dinucleotide AC in the mutation shared by the North American CHAC4 family and Mexican III-31 (3557_3558insAC) may have arisen by the same mechanism.
The shared geographical origin of most of the remaining probands with identical mutations suggests that these have been inherited from common ancestors. We could not perform haplotype analysis on certain probands in this study, so it was not possible to determine unambiguously whether they are related. However, haplotype analysis of the French Canadian families of II-42, III-6 and III-24, each harbouring a homozygous deletion (EX70_73del), revealed that the affected individuals were homozygous for a common haplotype spanning ~4 cM around VPS13A (haplotype A in Figure 6.2). French Canadian proband III-11 and her affected sister, both homozygous for the 4242+1G>T mutation, are homozygous for a different haplotype, as expected (haplotype B in Figure 6.2). However, two affected brothers in a different branch of the family are heterozygous for haplotypes A & B. This implies that these two brothers, who show classic symptoms of ChAc despite having only one copy of the ‘family’ mutation 4242+1G>T, may be heterozygous in addition for the EX70_73del mutation. Efforts are now being made to develop a junction PCR to amplify across the exon 70-73 deletion and unequivocally determine whether this is indeed the case.
Figure 6.2 - Haplotype analysis of French Canadian ChAc families. Haplotypes of affected individuals from four apparently unrelated families segregating ChAc are shown (proband labelled in white): the D9S1674 marker, which is internal to VPS13A, is indicated in red. Haplotype A is in pink, haplotype B is light yellow. Red asterisks denote where a homozygous deletion of exons 70-73 has been found by PCR analysis. Yellow asterisks denote where the splice-site mutation 4242+1G>T has been identified.

Key: *, heterozygous for mutation; **, homozygous for mutation.
The French Canadian population is of great interest to geneticists: it was founded by approximately 8500 settlers in the 17th century and has since increased in relative isolation to about six million people in Quebec province (Scriver 2001). Many Mendelian diseases occur at unusually high frequency in its subpopulations: for example, oculopharyngeal muscular dystrophy (OPMD, OMIM 164300) is found worldwide but is at greatest concentration in Southeastern Quebec. Analysis of haplotypes in 42 French Canadian OPMD families pointed to a single OPMD-associated haplotype in the patients, indicating a founder effect for this disorder (Brais et al. 1995). The discovery of four French Canadian families sharing a ChAc-associated haplotype indicates that ChAc may also show a founder effect in this population. Routine testing for the EX70_73del mutation may therefore be an option for rapid diagnosis of ChAc in French Canadian patients with relevant symptoms. Testing for the 4242+1G>T mutation, which was observed in a French proband (II-40), may also be worthwhile.

Interestingly, the EX70_73del mutation seems also to span the terminal two exons of GNA14, which encodes Gα subunit 14 (Figure 3.4). The α-subunits of G-proteins form heterotrimeric complexes with β- and γ-subunits; these complexes are involved in signal transduction from cell surface receptors to a variety of intracellular effectors. GNA14 is positioned just upstream of the gene encoding Gα subunit q, GNAQ: the two genes have probably arisen through tandem duplication (Wilkie et al. 1992). Although GNAQ is expressed in a wide range of tissues, GNA14 mRNA is virtually undetectable in adult tissue and appears to be restricted to foetal lung, kidney and liver, at levels far lower than GNAQ (Rubio et al. 1999). Preliminary investigations appear to show that the individuals who harbour the deletion of the 3' termini of
VPS13A and GNA14 have no obvious phenotypic differences from those who have mutations affecting VPS13A only. This would imply that either Gα14 function does not depend on the presence of its carboxy-terminal region, or that its absence is fully compensated for by, e.g., other Gα subunit proteins. Indeed, many biochemical activities ascribed to Gα14 have also been observed in other Gα proteins (e.g., (Nakamura et al. 1995; Maeda et al. 1996; Xu et al. 1998; Day et al. 2003)). However, detailed clinical analysis of individuals harbouring the GNA14 deletion may reveal subtle differences in phenotype that could give clues to Gα14’s physiological role.
## Table 6.1 - *VPS13A* mutations identified in two or more ChAc probands

<table>
<thead>
<tr>
<th>DNA change</th>
<th>Probands with mutation</th>
<th>Proband ancestry</th>
</tr>
</thead>
<tbody>
<tr>
<td>622OT</td>
<td>CHAC3 (ht)</td>
<td>Italian</td>
</tr>
<tr>
<td>1596-1G&gt;C</td>
<td>II-12 (ht)</td>
<td>Italian</td>
</tr>
<tr>
<td></td>
<td>II-27 (hm)</td>
<td>Japanese</td>
</tr>
<tr>
<td></td>
<td>II-39 (ht)</td>
<td>Danish</td>
</tr>
<tr>
<td>3557_3558insAC</td>
<td>III-29 (ht)</td>
<td>East German</td>
</tr>
<tr>
<td>4242+1G&gt;T</td>
<td>CHAC4 (hm)</td>
<td>North American</td>
</tr>
<tr>
<td>4355G&gt;C</td>
<td>III-31 (hm)</td>
<td>Mexican</td>
</tr>
<tr>
<td>4411C&gt;T</td>
<td>II-26 (ht)</td>
<td>Japanese</td>
</tr>
<tr>
<td>6059del</td>
<td>III-3 (hm)</td>
<td>Italian</td>
</tr>
<tr>
<td>6828del</td>
<td>II-37 (hm)</td>
<td>Israeli-Jewish</td>
</tr>
<tr>
<td></td>
<td>III-2 (hm)</td>
<td>Israeli-Jewish</td>
</tr>
<tr>
<td>9109G&gt;T</td>
<td>III-9 (ht)</td>
<td>Italian</td>
</tr>
<tr>
<td>EX70_EX73del</td>
<td>II-24 (hm)</td>
<td>French Canadian</td>
</tr>
<tr>
<td>9429_9432del</td>
<td>CHAC11 (ht)</td>
<td>English</td>
</tr>
<tr>
<td>9431_9432del</td>
<td>II-29 (ht)</td>
<td>British</td>
</tr>
<tr>
<td></td>
<td>III-18 (hm)</td>
<td>Irish</td>
</tr>
</tbody>
</table>

- Nucleotides and amino acids are numbered according to the cDNA sequence of *VPS13A* isoform 1A, reported by Rampoldi et al. (2001) as CHAC isoform A (GenBank accession no. NM_033305), with the adenosine of the initiation codon assigned position 1.
- Where we have no data on proband ancestry, the country in which the proband was examined is given in italics.

Key: ht, heterozygous; hm, homozygous
6.1.5 *Predicted effects of VPS13A mutations on function*

Fifty-eight of the 88 *VPS13A* mutations identified in this study (66%) are nonsense or insertion/deletion mutations that lead to PTCs and therefore predict absence or marked reduction of mutated *VPS13A* transcript via nonsense-mediated decay (Hentze and Kulozik 1999). This is in accordance with the recessive inheritance of choreo-acanthocytosis, where lack of chorein is the primary cause of the disease.

Nineteen out of the 88 *VPS13A* disease mutations described here (22%) are changes affecting splicing. A survey of point mutants that result in human genetic disease revealed that, on average, 10-15% were at splice junctions (Krawczak *et al.* 1992). A likely explanation for the relatively high percentage of splice-site mutations in ChAc is the large number of splice sites contained within *VPS13A*: its 143 splice sites presumably present quite a large target for mutation. In the absence of RNA from affected patients, it is difficult to predict exactly what effect these mutations have at the transcript level. However, as every splice-site mutation described here alters the respective exon/intron junction such that it does not score on the SpliceView program, it is likely that normal splicing will be markedly reduced in each case.

Only five of the *VPS13A* mutations that are predicted to cause disease in this study (5/88, 6%) are missense mutations (shown in Figure 6.3). They are non-conservative substitutions and were not found in at least 100 control chromosomes (384 control chromosomes were tested for Wave II mutations), which suggests that they do not represent benign polymorphisms. The substitution I90K occurs within the 190-aa amino-terminal region that is highly conserved among VPS13 homologues. The position is occupied by a hydrophobic amino acid in the chromosome 15 parologue of
VPS13A (VPS13C), as well as in VPS13A homologues in *Drosophila melanogaster*, *Caenorhabditis elegans*, *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*. The introduction of a positively charged residue here is likely to disrupt chorein structure and/or function.

The site of substitution A1095P is moderately well conserved, being occupied by a hydrophobic amino acid in all homologues described above. The introduction of the cyclic structure of proline here is likely to profoundly influence protein architecture, as it is often found in the bends of folded protein chains (Stryer 1995).

Although the identity of the amino acid is not so well conserved at the S1452P site, the introduction of a proline between the well-conserved hydrophobic amino acid at 1451 and leucine at 1453 is also likely to have steric consequences - as discussed below (section 6.2.2), this mutation leads to markedly reduced chorein levels. Intriguingly, the equivalent leucine residue is mutated to arginine in the VPS13B protein of a patient with Cohen syndrome, and is the sole missense mutation observed so far (L2193R, in (Kolehmainen *et al.* 2003)). Perhaps the precise architecture of this region is crucial to maintain protein integrity in VPS13 homologues.
Chapter 6: Discussion

The tryptophan site of the W2460R mutation is occupied by an aromatic or large
aliphatic amino acid in all VPS13A orthologues studied: its replacement for a
positively charged amino acid is therefore likely to have a deleterious effect.

Substitution Y2721C is in a locally well-conserved region – the introduction of the
reactive sulphhydryl group in cysteine may significantly influence the chemistry of this
area. The fact that just five out of 12 amino-acid variants reported here (Tables 3.1,
3.2 & 3.4) are believed to be pathogenic is perhaps an indication that much of chorein
is tolerant to substitution.
Figure 6.3 - Alignment of VPS13A missense mutations. Alignment of the amino-acid sequence surrounding the five missense mutations identified in this study. Sequences from chorein and homologues were aligned and coloured using the T-Coffee (Notredame et al., 2000) and CHROMA (Goodstadt et al., 2001) programs, respectively. The missense mutation site is arrowed. The equivalent position of the missense mutation L2193R reported in VPS13B is asterisked.

Key: Hs, Homo sapiens; Dm, Drosophila melanogaster; Ce, Caenorhabditis elegans; Sc, Saccharomyces cerevisiae; Sp, Schizosaccharomyces pombe.
6.2 Analysis of \textit{VPS13A} expression

6.2.1 \textit{Distribution of VPS13A mRNA and protein}
Northern and RT-PCR analysis revealed that \textit{VPS13A} mRNA is ubiquitously expressed (Figure 4.1). This is consistent with the fact that the 5'UTR of \textit{VPS13A} contains a CpG island, a feature characteristic of 'housekeeping' genes, and with the ubiquitous expression of its three human gene homologues (Velayos-Baeza \textit{et al.} 2004). Development of a polyclonal antiserum against chorein revealed that chorein is expressed in cell lines derived from a wide variety of human tissues, as well as primary skin fibroblasts and erythrocytes (Figures 4.7 - 4.9), which lends further support to the ubiquitous expression of this gene. Although the \textit{VPS13A} mRNA signal is weak (Figure 4.1), and experiments with the yeast homologue Vps13p indicate that its cell copy number is very low (Brickner and Fuller 1997), we found that endogenous chorein is expressed at a level sufficient to be detected by immunoblot assays of cell lysates.

6.2.2 \textit{Expression of mutant chorein in ChAc cell lines}
We have demonstrated that five classes of \textit{VPS13A} mutations (nonsense, frameshift, splice-site, missense, exon deletion) can lead to absence or marked reduction of chorein expression in ChAc patients (Figures 4.5, 4.7- 4.9). The affected siblings in family CHAC6 show very little expression of chorein (Figure 4.8a), implying that the missense mutation S1452P must adversely affect stability at the transcript or protein level. The fact that this missense mutation simply alters protein dosage is consistent with the theory that absence of chorein, rather than a functional defect of chorein, causes ChAc in the majority of patients. The affected son in CHAC11 is heterozygous
Chapter 6: Discussion

for an exon 72 mutation that should only affect the longer isoform of chorein, isoform 1A: isoform 1B expression (and that of other isoforms that don’t include exon 72) should in theory be unaffected. The fact that there is very little expression of chorein in this patient (Figure 4.8a) implies that isoform 1B is expressed at much lower levels than isoform 1A. As mentioned previously (section 6.1.3), we have demonstrated that isoform 1B and others cannot compensate for absence of chorein isoform 1A in chorea-acanthocytosis. It is therefore possible that exons 70-73 encode one or more protein domains that are essential for some functions of chorein. The relatively poor expression of isoforms lacking these exons, as demonstrated here, may provide an alternative reason for their inability to substitute for full-length chorein in this disorder.

6.2.3 Chorein detection for diagnosis of ChAc

We have shown that chorein can be detected in association with the erythrocyte membrane (Figure 4.9). It is still possible to detect chorein from blood that has been frozen for several months, with little apparent degradation of chorein in the samples. This has obvious implications for diagnosis of ChAc. At present, if the VPS13A gene has not been screened, ChAc can only be diagnosed by excluding other clinically similar disorders. Unlike in McLeod syndrome and abetalipoproteinemia (see sections 1.2 and 1.3), no simple biochemical test is at present possible. In contrast to Huntington’s disease (Kremer et al. 1994), no single mutation is responsible for the majority of cases of ChAc. Nor do the mutations cluster to a particular area of VPS13A: so far, ChAc disease mutations have been identified in 51/73 exons, implying that all regions of the gene must be screened for efficient mutation detection. As VPS13A is a large gene with many exons, screening is costly, labour-intensive and
time-consuming. Western blotting of patient erythrocyte membranes could perhaps
give an early indication of the disorder before a precise diagnosis is given by a
VPS13A gene screen.

Although chorein was absent or markedly reduced in all patient erythrocytes analysed
so far, one cannot exclude ChAc as a diagnosis if chorein is present in a sample. It is
possible that some patients harbour missense mutations that do not affect chorein
dosage but produce functionally defective chorein nevertheless. Also, some mutations
may allow almost normal expression levels of mutant chorein lacking only a few
exons, which might not be resolved from wild-type chorein due to its large size.
However, as missense changes comprise less than 6% of VPS13A mutations described
so far (this study and (Ueno et al. 2001), and the majority of transcripts containing
premature stop codons are believed to be rapidly degraded (Hentze and Kulozik
1999), we anticipate that this would not be a significant problem. In any case, if
chorein is absent in a patient sample, the most likely diagnosis will be chorea-
acanthocytosis.
6.3 Identification of potential chorein-interacting proteins

With only data from VPS13A homologues to draw on, it is clear that determination of chorein function is in its infancy. We used the yeast two-hybrid approach to search for proteins that may interact with chorein, in an attempt to infer the function of chorein from the function of its binding partners, and/or to dissect the pathological pathways that lead from chorein absence to development of ChAc. Although six ~600-aa regions of chorein were used as baits, and nearly $2.3 \times 10^7$ yeast transformants were screened, only seven independent clones, representing three distinct proteins, were identified. All were weak or very weak interactors with a single bait, chacY2 (Figure 5.3).

6.3.1 Limitations of the assay

It is likely that some physiological interactors with chorein have been missed in this screen. This could have been due to the bait selection. Due to its size, a Y2H screen using full-length chorein was unlikely to be successful, so baits comprising portions of the protein were designed. Logical bait design should take account of known domains of the protein, and of localisation signals that could interfere with nuclear localisation of the bait. However, due to limited structural information for chorein, this was not possible; and so functional domains of the protein may have been destroyed or mislocalised by bait selection. The recent discovery of the human chorein homologues on chromosomes 1, 8 and 15 has permitted further computer analysis of domains in chorein and some new conserved regions, such as 'domain of unknown function' 1162 and the R domains (as shown in Figure 6.4) have been identified. These regions could prove fruitful in future assays of chorein interactions.
Interactions may also have been missed due to bait sequestration by interaction with endogenous yeast orthologues of its physiological partners, or bait misfolding induced by expression as a fusion with GAL4 DB. In addition, post-translational modifications of the bait may be required for some interactions (Toby and Golemis 2001). Even if bait proteins were fully available, correctly folded and modified, some chorein-interactors may not have been represented in this cDNA library, due to low copy number in the brain. Screening of additional brain cDNA libraries, and libraries made from other tissues, may address this issue.

6.3.2 Potential chorein-binding partners

The potential chorein-interactors identified in this Y2H screen were TAB2, the putative protein PP2Cη and the p65 subunit of NF-κB (Figures 5.3-5.5). As three independent clones encoding TAB2 were identified, and TAB2 has already been characterised in other studies (e.g., Takaesu et al. 2000), we investigated the TAB2-chorein interaction further by co-immunoprecipitation studies in mammalian cells (Figures 5.6 & 5.7). The fact that some TAB2-T7 appears to be coprecipitated when chorein-FLAG is immunoprecipitated (and vice versa) implies that chorein and TAB2 may interact in vivo. The relatively low levels of lacZ activity induced by the chacY2-TAB2 interaction in the yeast two-hybrid studies and the fact that only a small amount of partner protein is coprecipitated in transfected cells implies that any in vivo interaction between chorein and TAB2 is likely to be transient and/or weak.
Figure 6.4 - Features of human VPS13 proteins. Schematic of VPS13A-D, splice form 1A. The amino (N) and carboxy (C) regions of high conservation are indicated in dark blue. The recent identification of this gene family and subsequent computer analysis of homologue alignments (Velayos-Baeza et al., 2004) has enabled elucidation of additional conserved regions: repeat regions 1-3 (R1-3) and carboxy-terminal region 2 (C2). Domain of unknown function 1162 (DUF1162) is a further region of homology shared between VPS13A, C and D. The chorein regions used for yeast two-hybrid analysis are indicated above the chorein plot. (After Velayos-Baeza et al., 2004)
We have not yet been able to determine the nature of the vesicular structures observed in immunofluorescent staining of chorein. Preliminary investigations of transiently transfected cells revealed that these structures did not colocalise with markers for the endoplasmic reticulum, Golgi, TGN, endosomes or other subcellular compartments (data not shown). Localisation studies in stably transfected cell lines are currently underway. Although transfected TAB2 and chorein did not colocalise (Figure 5.8), there is a possibility that the high expression levels of the recombinant proteins could result in mislocalisation and the true subcellular localisation of TAB2 is elsewhere. It would be interesting to repeat the immunofluorescent studies using an antibody for detection of endogenous TAB2, e.g., in neuronal cell lines, to resolve this issue. Subcellular fractionation studies indicate that IL-1 stimulation alters the distribution of TAB2 (Jiang et al. 2002); this implies that future chorein-TAB2 immunolocalisation studies could also benefit from different cytokine treatment conditions.

6.3.3 Possible relationships between chorein-interacting proteins

Intriguingly, experiments on TAB2, the type 2C protein phosphatases and NF-κB have illuminated potential relationships between these proteins. Some of the evidence linking them to each other and the same or similar complex signalling pathways is outlined below.

TAB1 & TAB2 were originally identified in a yeast two-hybrid screen as binding partners for transforming growth factor β (TGFβ)-activated kinase 1 (TAK1), a protein involved in the mitogen-activated protein (MAP) kinase cascade (Shibuya et al. 1996). TAK1 is a MAP kinase kinase kinase (MAPKKK) that is involved in
signalling mediated by interleukin-1 (IL-1), TGFβ and stress-related signals (Yamaguchi et al. 1995; Shirakabe et al. 1997). Jiang et al. (2002) proposed the following model of IL-1 signalling. Components from the stimulated IL receptor complex dissociate and interact with a complex of TAK1, TAB1 & TAB2 on the membrane. This leads to phosphorylation of TAK1 & TAB2 and translocation of the complex to the cytosol, where TAK1 is activated. TAK1 is implicated in the activation of IκB kinase (IKK) (Wang et al. 2001), which leads to NF-κB activation, as discussed below.

The PP2C proteins are the defining members of the PPM family of serine/threonine phosphatases. One of their roles is to reverse protein kinase cascades that become activated because of stress signals, including environmental stimuli and proinflammatory cytokines (Tamura et al. 2002). As mentioned above, TAK1 is one of the kinases involved in the stress response. PP2Cβ-1 and PP2Cε have both been shown to dephosphorylate and inactivate TAK1 (Hanada et al. 2001; Li et al. 2003). No experimental evidence yet exists for the human putative protein PP2Cη, but the orthologue in mouse has recently been characterised. Mouse PP2Cη shows strongest expression in testis, but modest expression was also observed in lung, kidney and brain (Komaki et al. 2003). Intriguingly, aa 64-368 of TAB1 are annotated in the SWISSPROT database as containing a PP2C-like domain and TAB1 aa 164-232 share 56% similarity with aa 53-120 of PP2Cη isoform 2; however, no phosphatase activity has been reported for TAB1.
NF-κB in its inactive form is present in the cytosol as a three-subunit complex, which usually comprises the transcription factor dimer of p50 & p65 and the inhibitory subunit IκB. Activation is effected by IKK: it causes the phosphorylation and subsequent degradation of IκB, leaving the p65-p50 dimer free to translocate to the nucleus and activate transcription of a wide range of genes. The activation of NF-κB in brain tissue by a diverse range of signals, including tumour necrosis factor alpha, nerve growth factor, glutamate and intracellular hydrogen peroxide has been documented (Mattson and Camandola 2001). Studies suggest that NF-κB plays important roles in cellular responses to injury of the nervous system in both acute and chronic neurodegenerative conditions. The consequences of NF-κB activation are complex: generally, in neurons it induces the expression of antiapoptotic gene products and other neuroprotective proteins, while NF-κB activation in microglia leads to expression of genes whose products promote neuronal degeneration (Mattson and Camandola 2001). Interestingly, mice lacking the p50 subunit of NF-κB undergo increased striatal neuron degeneration and have increased motor dysfunction when exposed to the mitochondrial toxin 3-nitropropionic acid, in an animal model of Huntington’s disease (Yu et al. 2000).
6.4 Potential functions for chorein

Although it has been established that mutations in \textit{VPS13A} are the primary cause of chorea-acanthocytosis, it is not yet known how chorein functions or how its abolition brings about the neurological and erythrocyte changes that make up the ChAc phenotype. Clinical data is currently being gathered for an exhaustive comparison of genotype and phenotype. However, as most of the \textit{VPS13A} mutations identified so far result in a null allele, and substantial intrafamilial phenotypic variation has been observed (e.g., in Aasly \textit{et al.} 1999), significant genotype-phenotype correlation is not likely. It may not be possible, therefore, to identify what clinical consequences are seen when different parts of the protein are ablated. As chorein also has no identifiable domains that can give clues to its normal function(s), information has to be derived from studies on its homologues.

6.4.1 Chorein: a protein sorter?

As discussed previously (section 1.8.1), the chorein homologue that has been most studied is Vps13p in \textit{Saccharomyces cerevisiae} (Redding \textit{et al.} 1996; Brickner and Fuller 1997). Vps13p is involved in the intracellular trafficking of Kex2p and other proteins: if we extrapolate from this finding, we can speculate that chorein may play a similar role in humans. Proteins that could be sorted by chorein include furin-like proteases, mammalian homologues of Kex2p. Furin has been shown to cycle between the TGN, endosomes and the plasma membrane (Nakayama 1997) and is therefore a good candidate for protein trafficking. Cytoskeletal components may also be mislocalised. In McLeod syndrome, the destabilisation of the Kell-XK complex and
the consequent disruption of contacts with the underlying membrane skeleton might lead to the structural changes thought to be related to acanthocyte formation (Terada et al. 1999). It is known that mice lacking the erythrocyte membrane protein 4.1R exhibit neurobehavioural deficits and abnormal erythrocyte morphology (Walensky et al. 1998). Perhaps mislocalisation of a similar protein or proteins in the absence of functional chorein could cause the haematological and neurological abnormalities seen in ChAc.

Additional potential targets for chorein protein sorting have been provided by the Y2H screen performed in this study. Again, it is feasible that chorein may be involved in the trafficking of one or more of these proteins: e.g., in shuttling of NF-κB between nucleus and cytosol, or movement of TAB2 between membrane and cytosol. The weakness of the interaction between chorein and TAB2 in vitro (and of chacY2 and its prey constructs in the Y2H screen) may be a reflection of the necessarily transient nature of an interaction between a sorting protein and its cargo. As NF-κB, TAB2 and PP2Cs all have links with the pathways involved in the cellular response to metabolic and environmental stress, mislocalisation of any of these proteins could possibly lead to a perturbation in the stress response. Under conditions of metabolic-excitotoxic stress, neurons with inadequate or inappropriate stress responses may be more liable to apoptose, thus leading to the marked neurodegeneration seen in the latter stages of chorea-acanthocytosis.
6.4.2 Possible redundancy in function of VPS13 proteins

A pertinent question is why the symptoms of ChAc appear to be limited to neuronal, muscular and erythroid cells, given the apparently ubiquitous expression of chorein. As mentioned previously (section 1.8.2), we have recently identified a novel human gene family comprising VPS13A and three other genes on chromosomes 1, 8 and 15. The proteins encoded by these genes show significant similarity with yeast Vps13p, notably at the amino- and carboxy-terminal regions (Velayos-Baeza et al. 2004). It is possible that these family members share some functions in the cell and that other VPS13 proteins can compensate for lack of chorein in most tissues. Only certain cells where chorein plays a non-redundant role would therefore be vulnerable in ChAc patients.

6.4.3 Future analysis of chorein function

Analysis of the few missense mutations that have been identified in this study might give important clues as to which amino acids are critical for chorein function. For example, the Y2H bait chacY2 is the site of the missense mutation W2460R: although time constraints prohibited further analysis, it would be interesting to see what effect the substitution of this residue has on the interactions between chacY2 and its putative binding partners. However, as we have shown, at least one of the missense changes appears simply to have an effect on protein dosage (Figure 4.8a), implying that a functional study of this mutant protein may not be biologically relevant. Western blotting of patient cells with missense mutations could be used to exclude such ‘dosage effect’ substitutions from further analysis, enabling efforts to be concentrated
on substitutions that may have functional implications. Functional analysis of these mutants is now needed in order to elucidate the exact role that chorein plays, both in the pathogenesis of ChAc and within the context of normal brain processes.


143


Fitch CD, Dinning JS (1963) Vitamin E deficiency in the monkey. V. Estimated requirements and the influence of fat deficiency and antioxidants on the syndrome. *J Nutr* 79:69-78


References


145
References


References


Lepke S, Fasold H, Pring M, Passow H (1976) A study of the relationship between inhibition of anion exchange and binding to the red blood cell membrane of 4,4'-diisothiocyanato stilbene-2,2'-disulfonic acid (DIDS) and its dihydro derivative (H2DIDS). J Membr Biol 29:147-177


contribution to the knowledge of the functional pathology of the caudate nucleus. *Acta Neuropathol* (Berl) 86:651-658


Marsh WL, Taswell HF, Oyen R, Nichols ME, Vergara MS, Pineda AA (1975b) Kx antigen of the Kell system and its relationship to chronic granulomatous disease. Evidence that the Kx gene is X-linked. *Transfusion* 15:527


150


References


152


retinitis pigmentosa, acanthocytosis, steatorrhea, and an abnormality of lipid metabolism. *Trans Am Neurol Assoc* 86:49-53


References


Abetalipoproteinemia-like lipid profile and acanthocytosis in a young woman

Regulation of stress-activated protein kinase signaling pathways by protein

Tanaka M, Hirai S, Kondo S, et al. (1998) Cerebral hypoperfusion and
hypometabolism with altered striatal signal intensity in chorea-acanthocytosis:
a combined PET and MRI study. Mov Disord 13:100-107

Tarugi P, Lonardo A (1997) Heterozygous familial hypobetalipoproteinemia
associated with fatty liver. Am J Gastroenterol 92:1400-1402

membrane skeletons in chorea-acanthocytosis and McLeod syndrome revealed
by the quick-freezing and deep-etching method. Acta Haematol 101:25-31

Thomas JG, Baneyx F (1997) Divergent effects of chaperone overexpression and
ethanol supplementation on inclusion body formation in recombinant
Escherichia coli. Protein Expr Purif 11:289-296

Toby GG, Golemis EA (2001) Using the yeast interaction trap and other two-hybrid-
based approaches to study protein-protein interactions. Methods 24:201-217

Towfighi J (1981) Effects of chronic vitamin E deficiency on the nervous system of
the rat. Acta Neuropathol (Berl) 54:261-267

membrane in the hereditary neurological disease with chorea, areflexia and
acanthocytosis. J Neurol Sci 56:89-97

protein, chorein, is mutated in chorea-acanthocytosis. Nat Genet 28:121-122

Analysis of the human VPS13 gene family. Genomics (submitted)


Vita G, Serra S, Dattola R, et al. (1989) Peripheral neuropathy in amyotrophic choreo-
acanthocytosis. Ann Neurol 26:583-587


**Appendix 1: PCR primers used in this study**

### A1.1 Primers for amplification of VPS13A exons

<table>
<thead>
<tr>
<th>Exon</th>
<th>Forward primer(^a)</th>
<th>Reverse primer(^b)</th>
<th>Amplicon size (bp)</th>
<th>AT(^\circ)C(^c), [Mg] (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACCGCCTCCGTCTCTGCTGG</td>
<td>GAAGGCCCGAGAACGACAG</td>
<td>362</td>
<td>62, 2(^*)</td>
</tr>
<tr>
<td>2</td>
<td>CTGATTAGACAGGAATGAGG</td>
<td>TAAGTTGACAAACTGATCGAGG</td>
<td>212</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>3</td>
<td>GTGAGTGATGAAATTGTGAAGG</td>
<td>GCTGAGAAACATGAAACTAC</td>
<td>275</td>
<td>54, 3</td>
</tr>
<tr>
<td>4</td>
<td>AGAAATAGACGAACTAAATGAC</td>
<td>CAAAGTGGAGCAAATAGTAGAG</td>
<td>339</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>5</td>
<td>CAAAATTAGTTATTAGTAGAGG</td>
<td>GAAATATCTCTCACTATGCGAG</td>
<td>295</td>
<td>54, 2.5</td>
</tr>
<tr>
<td>6</td>
<td>ATTACAAGGACAGGATGACAC</td>
<td>CTTTTGGGTTTTAAAGTAGTG</td>
<td>312</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>7</td>
<td>GGAGGCGGCTGCACTTAGACAG</td>
<td>GACACCTATTTCTGATCTGGG</td>
<td>249</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>8</td>
<td>TGAAGGAGGACAGTGCTAGGGG</td>
<td>TTATAGTGAGCCACATACAGAG</td>
<td>288</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>9</td>
<td>TTTAATCTGTGGATAAGGGTCAG</td>
<td>GATAGTACGACACACTCTCC</td>
<td>342</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>10</td>
<td>TGCCATGATATATCTAAATGAC</td>
<td>TGCAGCTGATCTACTATCTCG</td>
<td>469</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>11</td>
<td>CCACACATGCTCTGCTAGCTG</td>
<td>TGCCTTTATATCAAAATATATCC</td>
<td>343</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>12</td>
<td>CCAAATGCTGAGTAATGTCAG</td>
<td>CTCAATAGCCTAAAGGACCACAG</td>
<td>360</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>13</td>
<td>TGTGTATAGTAGAGGACATG</td>
<td>GGAAGTAAAGTGGAGGATGAC</td>
<td>342</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>14</td>
<td>GAAGGCTCTGTTCTGGTTCCTC</td>
<td>ATCAAAAGCAATTTGACCAC</td>
<td>315</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>15</td>
<td>GCCTCAAGGAGTTAGCTCTACG</td>
<td>GACACCCAGTGGTATTACTAGC</td>
<td>274</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>16</td>
<td>TTTATATGAAACGATACTATGTC</td>
<td>TATATTTCAAGTTTATAAGGCAAG</td>
<td>207</td>
<td>58, 3</td>
</tr>
<tr>
<td>17</td>
<td>GGAGTCATTATGATTGCTCTTG</td>
<td>ACTAGATATGTTGACCTATGTC</td>
<td>360</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>18</td>
<td>ATTGATAGTCTTCTTGTGGG</td>
<td>CTATATTTCTCTTCATCTCTGT</td>
<td>434</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>19</td>
<td>AATTTATTTTTCCTTTGATTTG</td>
<td>CTAGTTGAGACATTCCTGGT</td>
<td>367</td>
<td>54, 2.5</td>
</tr>
<tr>
<td>20</td>
<td>TGTGAAGATTATCTGAGGTC</td>
<td>AACAAAAAGACCCCAAATG</td>
<td>335</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>21</td>
<td>TACTTGTGAGATTATCTTCTCC</td>
<td>ATCACTTCCACACAAATCACTAAG</td>
<td>292</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>22</td>
<td>AATGGAATGTGTGACTATGAGG</td>
<td>AGGAGAGCTGTTATTAGTGG</td>
<td>291</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>23</td>
<td>TAAAGGCCAAATGGTATAGAG</td>
<td>TGTTGACTTTCTATAATCTG</td>
<td>327</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>24</td>
<td>CCGTTGAATATAACATTTTGTG</td>
<td>CTCTCTCTGTGCTCTTTGAG</td>
<td>230</td>
<td>62, 2.5</td>
</tr>
<tr>
<td>25</td>
<td>AGAGAGGCTAGTGGTTTTAGT</td>
<td>GAGTTACAGAGATACCACAG</td>
<td>402</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>26</td>
<td>GTTTATGATGATGCTGGTG</td>
<td>TCTAGCTGAGTGGGCAACTATG</td>
<td>374</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>27</td>
<td>CATTATTTTTCATTTACCTTCC</td>
<td>TACAAAGAGTTATATGCTACGT</td>
<td>254</td>
<td>54, 2.5</td>
</tr>
<tr>
<td>28</td>
<td>TGATTGTTATAGTGGATCTCC</td>
<td>GCTGATTTTTCTCTGATTC</td>
<td>414</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>29</td>
<td>TCTAATCTAGGAAATGATAGCC</td>
<td>CCTGCCATATAATGAACTCTCC</td>
<td>444</td>
<td>54, 2.5</td>
</tr>
<tr>
<td>30</td>
<td>AGTAGGAGATATTTCATACCTTG</td>
<td>519</td>
<td>58, 2.5</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>CAGGTACAGTAGAAATGTTGCC</td>
<td>CTATGATCTACTGATACATG</td>
<td>398</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>32</td>
<td>AAATAGTTCTGGTCTACAGA</td>
<td>TCTCAGCTGAAACACTGATTC</td>
<td>522</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>33</td>
<td>GATATTTTCCTTTTATACCAAC</td>
<td>TCAATGATGCTTTACATGCCC</td>
<td>360</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>34</td>
<td>GTGTTAGAGAGTTGAAACAGTC</td>
<td>GAGGTGCTGAATAGTACG</td>
<td>473</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>35</td>
<td>GCCAAAGATTAGTTGATTTATAC</td>
<td>TTTTCAGAGGGACAGAATAC</td>
<td>412</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>36</td>
<td>GCCCTCTCTTAGGCTCTTGG</td>
<td>TACTAGCTGACTGAATGAG</td>
<td>445</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>37</td>
<td>TCTCCAGGAGCATGCTCTTC</td>
<td>AAGGCTAACAGAGTAAAGCC</td>
<td>454</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>38</td>
<td>GGTGGTATAGTGGATTTTGGG</td>
<td>TTACAGACAGGTTGAGGAC</td>
<td>477</td>
<td>58, 3</td>
</tr>
<tr>
<td>39</td>
<td>AGATATTTAGTTTCTGATGCTC</td>
<td>CACACTATTTTTTGTGACATAAC</td>
<td>384</td>
<td>54, 2.5</td>
</tr>
<tr>
<td>40</td>
<td>TGACTGATGATGTTGAAAAGTG</td>
<td>TACAGAATTCTGAAATATTC</td>
<td>508</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>41</td>
<td>TACTGGAATGGGCTATGACTG</td>
<td>TGATGAGCAACCTGAACTTGG</td>
<td>488</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>Exon</td>
<td>Forward primer*</td>
<td>Reverse primer*</td>
<td>Amplicon size (bp)</td>
<td>AT (°C)*</td>
</tr>
<tr>
<td>------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>----------</td>
</tr>
<tr>
<td>43</td>
<td>CCTTCTAATGTTGATTTGAGG</td>
<td>ATGACAAAAGGAACAGATATAAGG</td>
<td>360</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>44</td>
<td>ACTTGTCTGATGTACTGCTGTTGAGG</td>
<td>CCTTCTAATGTTGATTTGAGG</td>
<td>474</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>45</td>
<td>TGCTGCTCTCTTAAAATATCCTCC</td>
<td>ACTGTTACGCTTATATTTTCTTCC</td>
<td>433</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>46</td>
<td>GCCCTAGTATTTCTCTCTTCTCAGTACCC</td>
<td>CACTGAAACCTAAAATACGTGTTAG</td>
<td>337</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>47</td>
<td>ACTACATATATATGAAAGGAGG</td>
<td>CTGATTACCTATTCCTTACTAG</td>
<td>523</td>
<td>54, 2</td>
</tr>
<tr>
<td>48</td>
<td>GACCTGTTCTGATGATTATTTC</td>
<td>TTAACAAAGTTATATTCTACCAAG</td>
<td>631</td>
<td>54, 2</td>
</tr>
<tr>
<td>49</td>
<td>GAAACCTGATAGAATATAACTTGG</td>
<td>TATACCTCTCTCTCTTCTGACCACACTG</td>
<td>528</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>50</td>
<td>CTGAAATTGAACTACTGAGTGGAGG</td>
<td>GGGGAGATTTAAACATTTAGCTTGG</td>
<td>382</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>51</td>
<td>CTAGACACCAGTCTTAAAACAGGG</td>
<td>ATCTTTGATCTCTTCTCTTCC</td>
<td>417</td>
<td>58, 3</td>
</tr>
<tr>
<td>52</td>
<td>GAAATTTTTTCGCTAGTCTCACC</td>
<td>TCGAGATCTCTAACAATACCT</td>
<td>321</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>53</td>
<td>CAGAGCAACGATACGCTCCTCC</td>
<td>AAAGGCTGCTGGGACACCC</td>
<td>283</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>54</td>
<td>CCATCAACCCTCATGAGTAGG</td>
<td>AGCCATACACATATACATG</td>
<td>491</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>55</td>
<td>GCCTACCTTGAATGAGGCTTG</td>
<td>TATACAAAGGTACTATGTACAGC</td>
<td>410</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>56</td>
<td>TGCTGAATTTTATGCGACTGAGG</td>
<td>GCTAGCACAACATTTAGAACTG</td>
<td>411</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>57</td>
<td>GTCTTTGCTACCCGCTAGAC</td>
<td>GTGTAAAGATCTTGGTGGTCC</td>
<td>350</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>58</td>
<td>GTAGTGAATTTGCGAAATACCTCAG</td>
<td>TGCTTTAATCTTGGCTCTTGC</td>
<td>319</td>
<td>54, 2</td>
</tr>
<tr>
<td>59</td>
<td>TTTGTGTATATAGGGTCTGTTCC</td>
<td>AAGGGTATGACACGAGAGAG</td>
<td>412</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>60</td>
<td>CCTCTGAGGCAACTATACAGC</td>
<td>GCTTATTGGCTCCTCAGTAC</td>
<td>435</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>61</td>
<td>TACGGATTTTTTTGACAATCC</td>
<td>GCTCATGACTGCCAAGCTAC</td>
<td>393</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>62</td>
<td>ACTGAGAGGAAAGGTTGAGG</td>
<td>ATCTTGAGGAAAATACCACATGC</td>
<td>338</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>63</td>
<td>TGCCATTAAGGTTTGGCTTG</td>
<td>CCTTCTCAGATTGTGGCTG</td>
<td>373</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>64</td>
<td>AGTCAATATACAATACCTCACA</td>
<td>ACCACCTGCTCTTAAATATCC</td>
<td>498</td>
<td>54, 2.5</td>
</tr>
<tr>
<td>65</td>
<td>TACCTAGAGAGCTCTGTAATTCC</td>
<td>ATGGCCTGAAAGCCAAGGACAG</td>
<td>473</td>
<td>56, 2</td>
</tr>
<tr>
<td>66</td>
<td>GACCTAAATCTCTCCTATAGTTCTG</td>
<td>TTAAGGCAAAATAGATAAAGGCC</td>
<td>416</td>
<td>54, 2</td>
</tr>
<tr>
<td>67</td>
<td>TAAATACAGCTGACTGATAGGC</td>
<td>TATGGGCATTATGGTATAAGTC</td>
<td>317</td>
<td>54, 2.5</td>
</tr>
<tr>
<td>68</td>
<td>GTTGGCCCATTGGAATATGGG</td>
<td>GAGAGCACCAAGATACGTGTTG</td>
<td>309</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>69</td>
<td>TTTTCATTTGCTAGTCTTCG</td>
<td>AAAGCGCTGGTAGATAGCCAG</td>
<td>366</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>70</td>
<td>AGGGCAGCATTTAGGCGTACAC</td>
<td>AGCTGGTGTTGAAATATTACCTG</td>
<td>349</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>71</td>
<td>ATGGGCTGAGGCTATACAC</td>
<td>AGCATCCAGATACAGGTTAC</td>
<td>294</td>
<td>54, 2.5</td>
</tr>
</tbody>
</table>

* When exons were amplified as a single amplicon, the forward primer is indicated on the row of the exon most 5', the reverse primer on the row of the exon most 3'.

b: A ‘touchdown’ program was used to amplify exons for DHPLC analysis (see section 2.2.2)

Key: AT, annealing temperature; [Mg], MgCl₂ concentration; * addition of deaza dGTP (see section 2.2.2)
## Appendix 1

### A1.2 Primers for amplification of VPS13A cDNA

<table>
<thead>
<tr>
<th>Region amplified (nt)</th>
<th>Primer name (direction)</th>
<th>Primer sequence</th>
<th>Amplicon size (bp)</th>
<th>AT (°C), [Mg] (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2193-2755</td>
<td>KIAArt13 (F)</td>
<td>AAACCTCAGTGTATCTACCCAGC</td>
<td>563</td>
<td>60, 2</td>
</tr>
<tr>
<td></td>
<td>KIAArt11 (R)</td>
<td>TCTCAATTTTCTGCCACCCAAATCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>516-1543</td>
<td>KIAArt22 (F)</td>
<td>GCTGTCATTGGTATTTCCCTTC</td>
<td>1028</td>
<td>58, 2.5</td>
</tr>
<tr>
<td></td>
<td>KIAArt15 (R)</td>
<td>TTACAATATTCCAGCTGCAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8871-9580</td>
<td>222006fp6 (F)</td>
<td>AGAAGGCCATCAGCTGGAGG</td>
<td>710</td>
<td>60, 2</td>
</tr>
<tr>
<td></td>
<td>BCD3239fp1 (R)</td>
<td>GTAGTCTAGGAAGCTGTAATCAG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Nucleotides are numbered according to the cDNA sequence of CHAC isoform A reported by Rampoldi et al. (2001) (Genbank accession no. NM_033305), with the adenosine of the initiation codon assigned position 1.

Key: F, forward; R, reverse; AT, annealing temperature; [Mg], MgCl$_2$ concentration

### A1.3 Primers for amplification of Chromosome 9 microsatellite markers

<table>
<thead>
<tr>
<th>Marker</th>
<th>Primer sequence</th>
<th>Label*</th>
<th>Amplicon size range (bp)</th>
<th>AT (°C), [Mg] (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GATA89a11</td>
<td>F: GCTTCTGAAAGCTTCTAGTTACC</td>
<td>TET</td>
<td>186-210</td>
<td>55, 2</td>
</tr>
<tr>
<td></td>
<td>R: AAATAGTAATGCGATTGCTAGGG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D9S1674</td>
<td>F: GCCTACCTCTGTAGACTGAC</td>
<td>FAM</td>
<td>216-232</td>
<td>53, 1.5</td>
</tr>
<tr>
<td></td>
<td>R: TTAGGAAAGCTGACCACTTCAAA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATA89c08</td>
<td>F: AGCCCAAGCTTTTAACTCATT</td>
<td>FAM</td>
<td>342-366</td>
<td>58, 2</td>
</tr>
<tr>
<td></td>
<td>R: GCTGCTCTGCTAAACATGCG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D9S153</td>
<td>F: TTAATGACAAGCGCAAATGGACTA</td>
<td>FAM</td>
<td>143-155</td>
<td>58, 1.5</td>
</tr>
<tr>
<td></td>
<td>R: GCAGATATGTTGCCAAAAACTCA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D9S1780</td>
<td>F: ACATTTGTCTTTATCCTGCATGA</td>
<td>TET</td>
<td>241-259</td>
<td>55, 1.5</td>
</tr>
<tr>
<td></td>
<td>R: CGCTTCTAAAGATAGCTGCTCA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D9S1867</td>
<td>F: GTGAACTGAGCTACCAAGCC</td>
<td>HEX</td>
<td>172-210</td>
<td>55, 2</td>
</tr>
<tr>
<td></td>
<td>R: ATCCGGCAGGGCTTCACACA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Fluorescent label was attached to forward primer only

Key: F, forward primer; R, reverse primer; AT, annealing temperature; [Mg], MgCl$_2$ concentration;
## A1.4 Primers for preparation of plasmid constructs

| Construct name | Primer name | Primer sequence | Amplicon size (bp) | Notes
|----------------|-------------|-----------------|-------------------|---
| pGEX-RT1       | XhoRT1F     | CCCTCGAGCTCTCTCTGGGATCTTGGAAGGCAATCTGGAAAGG | 918 | Xho site
|                | XhoRT1R     | CCCTCGAGATGGGGAAGGAGGCAATCTGGAAAGGCAATCTGGAAAGG | 918 | Xho site
| pGEX-RT2       | CHACRT1 (F) | CCCTCGAGCTCTCTCTGGGATCTTGGAAGGCAATCTGGAAAGGCAATCTGGAAAGG | 955 | Xho site
|                | XhoRT2R     | CCCTCGAGCTCTCTCTGGGATCTTGGAAGGCAATCTGGAAAGGCAATCTGGAAAGG | 955 | Xho site
| pGEX-RT1       | XhoRT3F     | CCCTCGAGCTCTCTCTGGGATCTTGGAAGGCAATCTGGAAAGGCAATCTGGAAAGG | 916 | Xho site
|                | XhoRT3R     | CCCTCGAGCTCTCTCTGGGATCTTGGAAGGCAATCTGGAAAGGCAATCTGGAAAGG | 916 | Xho site
| pENTR-chacY1   | attB1-chac1f| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1864 | attB site
|                | attB2-chac1r| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1864 | attB site
| pENTR-chacY2   | attB1-chac2f| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1853 | attB site
|                | attB2-chac2r| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1853 | attB site
| pENTR-chacY3   | attB1-chac3f| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1939 | attB site
|                | attB2-chac3r| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1939 | attB site
| pENTR-chacY4   | attB1-chac4f| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1870 | attB site
|                | attB2-chac4r| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1870 | attB site
| pENTR-chacY5   | attB1-chac5f| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1879 | attB site
|                | attB2-chac5r| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1879 | attB site
| pENTR-chacY6   | attB1-chac6f| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1876 | attB site
|                | attB2-chac6r| GGGAACACAGTGTACAACAAAAGGCAACCTCTATGGTTGAGTCAGG | 1876 | attB site
| pcD4chac2      | Kpn-chac2F  | CGGATCCGGGACAGTGCTGCTACACTCTCAATAGCTGCTCGG | 1814 | Kpn site
|                | Xho-chac2R  | CTCTCGAGCTCTCTCTGGGATCTTGGAAGGCAATCTGGAAAGG | 1624 | Xho site

* Restriction enzyme sites or recombination sites introduced by the primers are shown in bold. * PCR were all performed with annealing temperature of 60°C and magnesium acetate concentration of 1 mM, using rTth enzyme (section 2.2.3)

Key: F, forward; R, reverse; AT, annealing temperature; [Mg], MgCl2 concentration
A1.5 Primers to introduce restriction sites for *VPS13A* RFLP analysis

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
<th>Amplicon size (bp)</th>
<th>AT (°C), [Mg] (mM)</th>
<th>Enzyme site created</th>
<th>Base change analysed</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKex18RDF</td>
<td>GTCTTCATTGGAGATGCACATTGCCAC</td>
<td>188</td>
<td>58, 2.5</td>
<td><em>MwoI</em></td>
<td>1695C&gt;G</td>
</tr>
<tr>
<td>AKex18RDr</td>
<td>CATCAAATACCAACAACACACAGAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKex31RDF</td>
<td>AAATTTGCACAAAAAACAAAATTAAAAG</td>
<td>148</td>
<td>50, 2.5</td>
<td><em>BstXI</em></td>
<td>3283G&gt;C</td>
</tr>
<tr>
<td>AKex31RDr</td>
<td>GAACTTAAACAGATATTCTCATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKex38RDF</td>
<td>AAAAGACATGATGGATATAGTTG</td>
<td>166</td>
<td>58, 2.5</td>
<td><em>BglI</em></td>
<td>4469G&gt;A</td>
</tr>
<tr>
<td>AKex38RDr</td>
<td>TGCAACACTGTTCTACACAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKex54RDF</td>
<td>TGCCATATGATGGGAAAAAGTC</td>
<td>135</td>
<td>58, 2.5</td>
<td><em>HinDI</em></td>
<td>7457T&gt;C</td>
</tr>
<tr>
<td>AKex54RDr</td>
<td>GCTCTGCTACTCTGTCTTCTCAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: AT, annealing temperature; [Mg], MgCl₂ concentration

A1.6 Primers for amplification of *GNA14* exons

<table>
<thead>
<tr>
<th>Exon</th>
<th>Primer name</th>
<th>Primer sequence</th>
<th>Amplicon size (bp)</th>
<th>AT (°C), [Mg] (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>hG145'UTRfp2</td>
<td>CCGTCTTTTTCTTCCACACCCCTG</td>
<td>185</td>
<td>58, 2.5</td>
</tr>
<tr>
<td></td>
<td>hG14int1rp2</td>
<td>CCTCCGAGGCTGGTAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>hG14int1fp2</td>
<td>TCCCTCCAAATCCACACACCTCAG</td>
<td>415</td>
<td>58, 2.5</td>
</tr>
<tr>
<td></td>
<td>hG14int2rp2</td>
<td>TTCTCTGTGCTGGCAAAACACCTGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>hG14int2fp2</td>
<td>TCTCAAATCCACACAGGCAAGCGCCAGG</td>
<td>356</td>
<td>58, 2.5</td>
</tr>
<tr>
<td></td>
<td>hG14int3rp2</td>
<td>CTGCGACGGGATTCACGGGCGGG</td>
<td>353</td>
<td>58, 2.5</td>
</tr>
<tr>
<td>4</td>
<td>hG14int3fp2</td>
<td>CTGTGGGCTACCTGCTGCTGAN</td>
<td>353</td>
<td>58, 2.5</td>
</tr>
<tr>
<td></td>
<td>hG14int4rp2</td>
<td>GAAGGGCTACCTGCTGCTGGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>hG14int4fp2</td>
<td>TTACTGGGACTCCAGAGGCACTGT</td>
<td>353</td>
<td>54, 2.5</td>
</tr>
<tr>
<td></td>
<td>hG14int5rp2</td>
<td>GTAGCAAGACTCCTGCTACAATG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>hG14int5fp2</td>
<td>GACGGCGGCGGCGGGAGGGA</td>
<td>325</td>
<td>58, 2.5</td>
</tr>
<tr>
<td></td>
<td>hG14int6rp2</td>
<td>TAGACTGATGCCGGGAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>hG14int6fp2</td>
<td>GAGAGGACATGTGTTAAGCTCTGTG</td>
<td>342</td>
<td>58, 2.5</td>
</tr>
<tr>
<td></td>
<td>hG143'UTRrp2</td>
<td>TGGGGCTGGCTGGTGAGTCTG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: AT, annealing temperature; [Mg], MgCl₂ concentration
### A1.7 Primers for amplification and sequencing of XK exons

<table>
<thead>
<tr>
<th>Exon</th>
<th>Primer name (direction)*</th>
<th>Primer sequence</th>
<th>Amplicon size (bp)</th>
<th>AT (°C), [Mg] (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>XKbfp1 (F)</td>
<td>TGACGCGCGGAGATGAAATTC</td>
<td>432</td>
<td>61, 1.5</td>
</tr>
<tr>
<td></td>
<td>XKrp1 (R)</td>
<td>ATGATTCGTCGCTTCCTTGACT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XKseq1 (R)</td>
<td>AGATCAGGCTACAGGACGGCT</td>
<td></td>
<td>(Sequencing)</td>
</tr>
<tr>
<td>2</td>
<td>XKbfp2 (F)</td>
<td>AGGACTCATACTGAGACTGTC</td>
<td>625</td>
<td>61, 1.5</td>
</tr>
<tr>
<td></td>
<td>XKrp2 (R)</td>
<td>CCCCTGGCTCTAGATGCAAGGTTCATA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XKseq2.1 (F)</td>
<td>GTTCAAGATAGCGCAAGTCAGGCT</td>
<td></td>
<td>(Sequencing)</td>
</tr>
<tr>
<td></td>
<td>XKseq2.2 (R)</td>
<td>GAGCATAGTAACTGACAGAGTTCTACAAC</td>
<td></td>
<td>(Sequencing)</td>
</tr>
<tr>
<td>3</td>
<td>XKbfp3 (F)</td>
<td>AACTGGAAGTGCTAGGCTGTGCACAT</td>
<td>1217</td>
<td>65, 1.5</td>
</tr>
<tr>
<td></td>
<td>XKrp3 (R)</td>
<td>GGCCAGTATAATGGCTAGAGAACAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XKseq3.1 (F)</td>
<td>TATCCCTGTTGCTCCATGATG</td>
<td></td>
<td>(Sequencing)</td>
</tr>
<tr>
<td></td>
<td>XK3.1bis (F)</td>
<td>TGACAGAGTGCTGCTGCAGTAG</td>
<td></td>
<td>(Sequencing)</td>
</tr>
<tr>
<td></td>
<td>XKseq3.2 (R)</td>
<td>ACCAGGACTACAATCGGATGCGC</td>
<td></td>
<td>(Sequencing)</td>
</tr>
<tr>
<td></td>
<td>XKseq3.3 (R)</td>
<td>ATCTCGGACTGACTGTAGATCTTC</td>
<td></td>
<td>(Sequencing)</td>
</tr>
<tr>
<td></td>
<td>XKseq3.4 (F)</td>
<td>GAACTAGAGAGAGGCCCTACAGTAG</td>
<td></td>
<td>(Sequencing)</td>
</tr>
</tbody>
</table>

* Primers used for PCR amplification are shown in bold.
Key: AT, annealing temperature; [Mg], MgCl₂ concentration
Appendix 2: GenBank accession numbers for sequences used in this study

This appendix lists the accession numbers for cDNA and protein sequences deposited at the GenBank database at the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/).

A2.1 Accession numbers for cDNA sequences

H sapiens FLJ39362 AK096681
H sapiens FLJ32332 NM_144641
H sapiens NF-κB p65 subunit M62399
H sapiens TAB2 splice variant 2 NM_145342
H sapiens VPS13A (CHAC) splice variant 1A NM_033305
H sapiens VPS13A (CHAC) splice variant 1B NM_015186
## A2.2 Accession numbers for protein sequences

<table>
<thead>
<tr>
<th>Organism</th>
<th>Protein Name</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. thaliana</em></td>
<td>AT1g48090.1</td>
<td>NP_175242</td>
</tr>
<tr>
<td><em>A. thaliana</em></td>
<td>VPS13-like protein</td>
<td>CAC08248</td>
</tr>
<tr>
<td><em>A. thaliana</em></td>
<td>O23559</td>
<td>CAB80985</td>
</tr>
<tr>
<td><em>A. thaliana</em></td>
<td>O23557</td>
<td>CAB80983</td>
</tr>
<tr>
<td><em>C. elegans</em></td>
<td>T08G11.1</td>
<td>T24692</td>
</tr>
<tr>
<td><em>D. discoideum</em></td>
<td>TipC</td>
<td>AAC31916</td>
</tr>
<tr>
<td><em>D. melanogaster</em></td>
<td>CG2093</td>
<td>AAF59228</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>chorein variant 1A</td>
<td>NP_150648</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>chorein variant 1B</td>
<td>NP_056001</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>NF-κB p65 subunit</td>
<td>AAA36408</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>PP2Cη</td>
<td>Q96ML6</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>TAB1</td>
<td>Q15750</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>TAB2 variant 2</td>
<td>NP_663317</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>VPS13B variant 1A</td>
<td>CAE75584</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>VPS13C variant 1A</td>
<td>CAE75582</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>VPS13D variant 1A</td>
<td>CAE75586</td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>Vps13p/Soilp</td>
<td>NP_013060</td>
</tr>
<tr>
<td><em>S. pombe</em></td>
<td>T39553</td>
<td>T39553</td>
</tr>
<tr>
<td><em>S. pombe</em></td>
<td>VPS13 orthologue</td>
<td>NP_596800</td>
</tr>
</tbody>
</table>
Appendix 3: Publications arising from this study

Chorein detection for the diagnosis of chorea-acanthocytosis
Ann Neurol (in press)

Velayos-Baeza A, Vettori A, Copley RR, Dobson-Stone C, Monaco AP. 
Analysis of the human VPS13 gene family
Genomics (in press)

Chorea-acanthocytosis: Clinical and genetic findings in three families from the Arabian peninsula 

Mutational spectrum of the CHAC gene in patients with chorea-acanthocytosis. 

McLeod neuroacanthocytosis: genotype and phenotype. 

A conserved sorting-associated protein is mutant in chorea-acanthocytosis. 