European Journal of Human Genetics (2002) 10, 773–781

www.nature.com/ejhg

ARTICLE

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

Carol Dobson-Stone[1], Adrian Danek[2], Luca Rampoldi[3], Richard J Hardie[4], Richard M Chalmers[5], Nicholas W Wood[5], Saeed Bohlega[6], Maria Teresa Dotti[7], Antonio Federico[7], Masami Shizuka[8], Makoto Tanaka[8], Mitsunori Watanabe[8], Yoshio Ikeda[8], Mitchell Brin[9], Lev G Goldfarb[10], Barbara I Karp[10], Saidi Mohiddin[11], Lameh Fananapazir[11], Alexander Storch[12], Alan E Fryer[13], Paul Maddison[14], Igor Sibon[15], Paulo C Trevisol-Bittencourt[16], Carlos Singer[17], Ignacio Requena Caballero[18], Jan O Aasly[19], Klaus Schmierer[20], Reinhard Dengler[21], Lutz-Peter Hiersemenzel[22], Massimo Zeviani[23], Vardiella Meiner[24], Alexander Lossos[25], Sturla Johnson[26], Flavio C Mercado[27], Giuseppe Sorrentino[28], Nicolas Dupré[29], Guy A Rouleau[29], Jens Volkmann[30], Javier Arpa[31], Andrew Lees[32], Gilles Geraud[33], Sylvain Chouinard[34], Andrea Németh[1,35] and Anthony P Monaco*,[1]

[1]The Wellcome Trust Centre For Human Genetics, University of Oxford, UK; [2]Neurologische Klinik, Ludwig-Maximilians-Universität, Munich, Germany; [3]DIBIT, HSR Scientific Park, Human Molecular Genetics Unit 2A3, Milan, Italy; [4]St George's and Atkinson Morley's Hospitals, London, UK; [5]Department of Molecular Pathogenesis, Institute of Neurology, London, UK; [6]Department of Neurosciences, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia; [7]Neurometabolic Unit, Institute of Neurological Sciences, University of Siena, Italy; [8]Department of Neurology, Gunma University School of Medicine, Maebashi, Japan; [9]Mount Sinai School of Medicine, New York, NY and Allergan, Inc, Irvine, CA, USA; [10]National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, USA; [11]Inherited Cardiac Disease Section, Cardiovascular Branch, National Heart Lung and Blood Institute, Bethesda, USA; [12]Department of Neurology, University of Ulm, Ulm, Germany; [13]Department of Clinical Genetics, Royal Liverpool Children's Hospital (Alder Hey), Liverpool, UK; [14]Department of Neurology, Pinderfields Hospital, Wakefield, UK; [15]Federation of Clinical Neurosciences, Hôpital Pellegrin, Bordeaux, France; [16]Federal University of Santa Catarina, Brazil; [17]Department of Neurology, University of Miami School of Medicine, Miami, USA; [18]Sección de Neurología, Hospital de Conxo, Complexo Hospitalario Universitario, Santiago de Compostela, Spain; [19]Department of Neurology, University Hospital, Trondheim, Norway; [20]Klinik und Poliklinik für Neurologie, Charité, Humboldt-Universität, Berlin, Germany; [21]Department of Neurology, Medical School Hannover, Germany; [22]Psychiatrische Klinik Oberwil, Oberwil, Switzerland; [23]Istituto Nazionale Neurologico C. Besta, Milan, Italy; [24]Department of Human Genetics and; [25]Department of Neurology, Hadassah University Hospital, Jerusalem, Israel; [26]Akershus Universitetssykehus, Nordbyhagen, Norway; [27]Facultad de Medicina, Universidad de Buenos Aires, Argentina; [28]Faculty of Sciences of the Movement, University of Naples Parthenope, Naples, Italy; [29]Center for Research in Neurosciences, McGill University and the McGill University Health Center Research Institute, Montreal, Canada; [30]Neurologische Klinik, Christian-Albrechts-Universität zu Kiel, Kiel, Germany; [31]Servicio de Neurología, Hospital Universitario, La Paz, Madrid, Spain; [32]Reta Lila Weston Institute for Neurological Studies, Royal free and UCL School of Medicine, Windeyer Medical Institute, London, UK; [33]Department of Neurology, CHU Rangueil, Toulouse, France; [34]Unité des Troubles du Mouvement André-Barbeau, CHUM Hotel-Dieu, Montréal, Canada; [35]Department of Clinical Genetics, Churchill Hospital, Oxford, UK

***Correspondence: AP Monaco, The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Headington, Oxford, OX3 7BN, UK. Tel: +44 1865 287502; Fax: +44 1865 287650; E-mail; anthony.monaco@well.ox.ac.uk**
Received 16 April 2002; revised 28 June 2002; accepted 1 July 2002

Chorea-acanthocytosis (ChAc) is an autosomal recessive neurological disorder whose characteristic features include hyperkinetic movements and abnormal red blood cell morphology. Mutations in the *CHAC* gene on 9q21 were recently found to cause chorea-acanthocytosis. *CHAC* encodes a large, novel protein with a yeast homologue implicated in protein sorting. In this study, all 73 exons plus flanking intronic sequence in *CHAC* were screened for mutations by denaturing high-performance liquid chromatography in 43 probands with ChAc. We identified 57 different mutations, 54 of which have not previously been reported, in 39 probands. The novel mutations comprise 15 nonsense, 22 insertion/deletion, 15 splice-site and two missense mutations and are distributed throughout the *CHAC* gene. Three mutations were found in multiple families within this or our previous study. The preponderance of mutations that are predicted to cause absence of gene product is consistent with the recessive inheritance of this disease. The high proportion of splice-site mutations found is probably a reflection of the large number of exons that comprise the *CHAC* gene. The *CHAC* protein product, chorein, appears to have a certain tolerance to amino-acid substitutions since only two out of nine substitutions described here appear to be pathogenic.

European Journal of Human Genetics (2002) **10,** 773–781. doi:10.1038/sj.ejhg.5200866

Keywords: choreoacanthocytosis; neuroacanthocytosis; mutational spectrum; *CHAC*; chorein

Introduction

Chorea-acanthocytosis (ChAc, OMIM 200150) is an uncommon neurodegenerative disorder. First described more than 30 years ago by Levine[1] and Critchley,[2] the disease is mainly characterised by gradual onset of hyperkinetic movements and aberrant erythrocyte morphology.[3] Diseases that combine neurological features with acanthocytosis have been termed 'neuroacanthocytoses'. Other disorders in this group include abetalipoproteinemia (OMIM 200100), hypobetalipoproteinemia (OMIM 107730) and McLeod syndrome (OMIM 314850). ChAc is an autosomal recessive condition, although autosomal dominant transmission has been reported.[1, 4, 5] The onset of neurological symptoms is usually delayed until 25–45 years of age and the disease follows a progressive course.[3] There is as yet no effective long-term treatment for this disorder.

As well as chorea, which closely resembles that observed in Huntington's disease, patients can suffer from motor and vocal tics, dystonia, parkinsonism and ocular motor disorders.[3,6,7] Orofacial dyskinesia often leads to tongue and lip biting and dysphagia. Epileptic seizures and cognitive as well as psychopathological abnormalities are also seen.[3,8] Post-mortem studies reveal degeneration of the basal ganglia, which results in atrophy of the putamen and caudate nucleus. In contrast to Huntington's disease, however, the cerebral cortex is usually spared.[3]

The term acanthocytosis is used to describe the unusual spiky appearance of a proportion of red blood cells in the patients' blood. Reported percentages of acanthocytes in affected individuals' blood vary widely, but are usually in the 5–50% range.[3] As there is no standard clinical method for performing an acanthocyte count, it is not clear how much variability can be accounted for by differing laboratory procedures.[9] The broader significance and molecular basis of acanthocytosis is unclear.

We reported linkage of chorea-acanthocytosis in 11 families to a 6 centimorgan interval flanked by the markers *GATA89a11* and *D9S1843* on chromosome 9q21 in 1997.[10] We subsequently identified a novel gene in the ChAc critical region and found 16 different mutations in these same 11 families.[11] Another group who independently identified the gene reported one other mutation in three Japanese families.[12] The *CHAC* gene is organised in 73 exons spanning about 250 kb of the genome. We identified two main transcripts: transcript A comprises exons 1–68 and 70–73, encoding the 3174 amino acid protein chorein, whereas transcript B contains exons 1–69 only.

Chorein seems to belong to no known human gene family. Computer searches of its sequence do not identify any known structural motifs or domains. Even the number of putative transmembrane domains in chorein diverges widely between different prediction programs. Other than the assumption that chorein plays a similar role to its structural homologue in *Saccharomyces cerevisiae*,[13] little can be said of its potential function or how its absence leads to the chorea-acanthocytosis phenotype. As the first step towards addressing these issues and to obtain an idea of the mutational spectrum of this disorder, we conducted a large-scale screen of 43 unrelated probands with ChAc. We report 54 novel mutations in *CHAC* and discuss their composition and distribution within the gene.

Materials and methods

Patients

Forty-three probands from 15 countries were included in this study. Thirty-three probands were the only known affected member of their family; the other 10 came from families with multiple affected members. No families that have been reported to display autosomal dominant inheritance were included in this study. The clinical diagnosis

of ChAc was based on the presence of acanthocytes in patients with a movement disorder in whom McLeod syndrome had been excluded. The clinically very similar family of Johnson in whom acanthocytes had not been found[14] was also included. Peripheral blood was collected for DNA isolation after patients had given informed consent. Where possible, families were haplotyped and only those with results consistent to linkage to the *CHAC* locus were included in this study.

Mutation analysis

For 14 probands, genomic DNA was extracted from peripheral blood in Oxford using the Nucleon kit (Scotlab, Lanarckshire, UK). The remainder were extracted using a variety of procedures. All 73 translated *CHAC* exons plus flanking splice sites were amplified from the proband DNA. Primers used were mainly as published previously,[11] but some were redesigned for improved mutation analysis (primer sequences and PCR conditions listed in Supplementary Information). PCR amplification was carried out with standard reaction mixes containing a 9 : 1 ratio of AmpliTaq Gold™: *Pfu* Turbo DNA polymerases (PE-Applied Biosystems, Cheshire, UK; and Stratagene, Amsterdam, Holland, respectively). Thermocycling was carried out as follows: 95°C for 15 min; a touchdown of 14 cycles (95°C for 30 s; annealing temperature +7.5°C with −0.5°C per cycle for 30 s; 72°C for 30 s); 25 cycles of 95°C for 30 s, the annealing temperature for 30 s, 72°C for 30 s; 72°C for 7 min. Amplification products from probands were combined in a 3 : 1 ratio with the appropriate wild-type homozygous amplicon. This mixture was then annealed and analysed by denaturing high-performance liquid chromatography (DHPLC), as described previously.[15] PCR products showing a variant DHPLC pattern were purified using a QIAquick PCR purification kit (QIAGEN, Surrey, UK). These were then sequenced in both directions by means of the BigDye Terminator Cycle Sequencing Ready Reaction kit (PE-Applied Biosystems) and were run on an ABI PRISM 377 DNA Sequencer (PE-Applied Biosystems). For 11 exons (exons 2, 16, 29, 33, 35, 48, 49, 50, 66, 67 and 72), analysis by DHPLC was not appropriate, either because the amount of product was not sufficient or because the amplicon contained a high frequency polymorphism. These were therefore sequenced directly as above. The presence of each mutation was verified in the proband and any available family members by restriction digestion or sequencing (Table 1).

Restriction enzyme analysis of controls

The DNA of 192 control individuals (European Collection of Cell Cultures, Wiltshire, UK) was analysed to check for the presence of any missense or exonic splice-site mutations by restriction enzyme analysis (Table 1). Where the mutations did not provide a restriction fragment polymorphism, a diagnostic restriction site was introduced by PCR mutagenesis (primer sequences listed in Supplementary Information).

Haplotype analysis

Polymorphic markers *GATA89A11*, *D9S1674*, *GATA89C08*, *GGAT13B07*, *D9S153*, *AFMA101XD1*, *D9S1867* and *D9S1843* flanking the *CHAC* gene were used in order to define haplotypes and establish linkage to the *CHAC* locus. Amplification conditions were as described previously.[10]

Southern blot analysis

Where it was repeatedly not possible to amplify a certain exon in a proband, Southern blot analysis was performed on *Bam*HI, *Hin*dIII and *Eco*RI digests of the patient's genomic DNA according to standard procedures.[16]

Results

The entire coding and flanking intronic sequence in *CHAC* was screened for mutations by DHPLC in 43 unrelated patients with chorea-acanthocytosis. This revealed 57 different mutations likely to cause disease in 39 probands (Table 1). Only three of these mutations have been previously reported, in other families.[11] All available family members were screened for the relevant mutation(s) using the verification method outlined in Table 1; 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.

Nonsense mutations

Seventeen different base substitutions that result in premature termination codons (PTCs) were identified (Table 1). 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.

Splice-site mutations

Fifteen different mutations are predicted to result in altered *CHAC* mRNA splicing (Table 1). 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 *CHAC* mRNA, we consider its primary effect to be on splicing, since the new intron/exon junction formed (attcttt-agTATTTCAGTATTAT → attctttaTTAT) 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 CCGgtaata → CCAgtaata: although both trinucleotides CCG and CCA code for proline, this exon/intron junction no longer

Table 1 *CHAC* disease mutations

Location	*DNA change*[a]	*Protein change*[a]	*Proband no.*[c]	*Verification method*
Nonsense				
Exon 9	622C>T[b]	R208X[b]	12 (ht), 27* (hm), 39 (ht)	Sequencing
Exon 17	1549G>T	E517EX	3 (ht)	*Apo*I
Exon 18	1616C>G	S539X	18 (ht)	*Tfi*I
Exon 25	2593C>T	R865X	9 (hm)	*Taq*I
Exon 29	3109A>T	K1037X	8 (ht)	Sequencing
Exon 30	3157C>T	Q1053X	15* (hm)	Sequencing
Exon 34	3889C>T	R1297X	38 (ht)	Sequencing
Exon 37	4355C>G	S1452X	35 (ht)	Sequencing
Exon 37	4411C>T	R1471X	26 (ht)	*Sty*I
Exon 45	5920G>T	E1974X	5* (hm)	*Dra*I
Exon 46	6094C>T	R2032X	4 (ht)	*Bsr*I
Exon 48	6419C>G	S2140X	24 (ht)	*Ban*II
Exon 48	6494G>A	W2165X	10 (hm)	Sequencing
Exon 50	7005G>A	W2335X	36 (ht)	*Ple*I
Exon 56	7867C>T	R2623X	23 (ht)	Sequencing
Exon 68	9109C>T[b]	R3037X[b]	1 (ht)	*Tsp*45I
Exon 70	9219C>G	Y3073X	13 (ht)	Sequencing
Splice site				
Intron 11	883-1_892del	(SA: no score (wt=82))[d]	22 (ht)	Sequencing
Intron 17	1595+1G>A	(SD: no score (wt=86))	36 (ht)	Sequencing
Intron 17	1596-2A>C	(SA: no score (wt=89))	16 (ht)	Sequencing
Intron 17	1596-1G>C	(SA: no score (wt=89))	35 (ht)	Sequencing
Intron 21	2170+1G>A	(SD: no score (wt=85))	26 (ht)	*Hph*I
Intron 22	2288+2T>C	(SD: no score (wt=83))	2 (ht)	Sequencing
Intron 36	4242+1G>T	(SD: no score (wt=93))	40 (hm)	Sequencing
Intron 40	4956+1G>T	(SD: no score (wt=83))	12 (ht)	Sequencing
Exon 55	7806G>A[e]	(SD: no score (wt=76))	3 (ht)	*Hpa*II
Intron 61	8472-1G>C	(SA: no score (wt=88))	2 (ht)	Sequencing
Intron 65	8907+2T>A	(SD: no score (wt=75))	16 (ht)	Sequencing
Intron 70	9275+1G>A	(SD: no score (wt=80))	1 (ht)	*Mae*III
Intron 70	9276-2A>T	(SA: no score (wt=92))	41 (ht)	Sequencing
Intron 71	9399+2_+8del	(SD: no score (wt=92))	14 (ht)	Sequencing
Exon 72	9474G>A[e]	(SD: no score (wt=81))	8 (ht)	*Scr*FI
Missense				
Exon 31	3283G>C[e]	A1095P	21 (ht)	*Bst*XI
Exon 53	7378T>C[e]	W2460R	28 (hm)	*Aci*I
Small insertion/deletion				
Exon 13	994del	331WfsX9	7 (ht)	Sequencing
Exon 13	1125_1128del	373TfsX22	11 (ht)	Sequencing
Exon 14	1187_1188del	395VfsX0	17* (hm)	Sequencing
Exon 14	1208_1211del	402RfsX5	23 (ht)	Sequencing
Exon 20	2029_2031delins27	676GinsIYX	29 (ht)	*Psh*AI
Exon 34	3847del	1282VfsX6	33 (hm)	*Bsr*I
Exon 36	4216del	1405MfsX19	38 (ht)	*Sty*I
Exon 37	4346del	1448FfsX4	19 (ht)	*Mbo*II
Exon 39	4835del	1611LfsX29	21 (ht)	*Bsr*I
Exon 41	5253_5266del	1750RfsX0	13 (ht)	*Dra*I
Exon 46	6059del	2019SfsX8	37 (hm)	*Hph*I
Exon 47	6283del	2094YfsX9	7 (ht)	Sequencing
Exon 49	6804_6805insG	2268LfsX6	39 (ht)	*Ava*II
Exon 49	6828del	2276TfsX11	41 (ht)	Sequencing
Exon 53	7339_7340insT	2446FfsX4	20 (ht)	Sequencing
Exon 57	7985_7989del	2661FfsX5	32 (ht)	Sequencing
Exon 57	8007del	2668VfsX21	22 (ht)	*Hpa*I
Exon 61	8390del	2796GfsX1	14 (ht)	*Mn*I
Exon 70	9190del	3063QfsX16	24 (ht)	Sequencing
Exon 72	9429_9432del[b]	3142AfsX4[b]	34 (ht)	*Bsr*I
Exon 72	9431_9432del	3143RfsX5	29 (ht)	*Pfl*MI
Gross deletion				
Exon 23	EX23del	(unknown)	6* (hm)	Southern blot
Exons 70–73	EX70_EX73del	(unknown)	42 (hm)	Southern blot

[a]Nucleotides and amino acids are numbered according to the cDNA sequence of *CHAC* isoform a reported by Rampoldi *et al* ([11]; Genbank accession no. NM_033305), with the adenosine of the initiation codon assigned position 1. Mutations are described according to the nomenclature recommended by Dunnen and Antonarakis ([21] and http://www.dmd.nl/mutnomen.html). [b]Mutations that have been identified in another family previously[11]. [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 inputted 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. hm, homozygous; ht, heterozygous; SA, splice acceptor; SD, splice donor; wt, wild type.

conforms to the splice-donor consensus. The same is true of mutation 9474G>A in proband 8: the substitution AGGgtaaat → AGAgtaaat does not affect the coding for arginine 3158, but does abolish the splice-donor site. Both exonic splice-site mutations were found to be absent in 192 control individuals using the verification methods indicated in Table 1. Unfortunately, as RNA was not available from these families, it was not possible to determine the precise effect of these mutations on splicing. However, input of all these base changes into the SpliceView program (http://l25.itba.mi.cnr.it/~webgene/wwwspliceview.html) revealed that, in each case, the new junction sequence was too divergent from the consensus to score as a splice-site (Table 1).

Missense mutations

Only two different base changes resulting in a potentially pathogenic, non-conservative amino-acid substitution were identified (3283G>C & 7378T>C, Table 1). Neither of these missense mutations was found when screening 192 normal individuals using the verification methods outlined in Table 1. Mutation 3283C>G results in the substitution of alanine to proline at position 1095: this position is occupied by an aliphatic amino acid in putative chorein homologues T08G11.1, CG2093, VPS13 and KIAA1421 of *Caenorhabditis elegans*, *Drosophila melanogaster, Saccharomyces cerevisiae* and *Homo sapiens*, respectively. Mutation 7378T>C results in the substitution of tryptophan to arginine at position 2460, which is occupied by a large aliphatic or aromatic amino acid in the aforementioned homologues.

Insertion/deletion mutations

Twenty different insertions or deletions leading to a shift in the reading frame, introducing a PTC, were identified (Table 1). 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 6 and 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 three terminal exons of *CHAC* (Figure 1).

Exonic polymorphisms

In addition to the mutations listed above, 16 exonic variants were identified (Table 2). Nine of these represent silent mutations, in that the nucleotide substitutions do not alter the amino-acid coding. The remaining seven changes were concluded to be neutral variants, despite leading to amino-

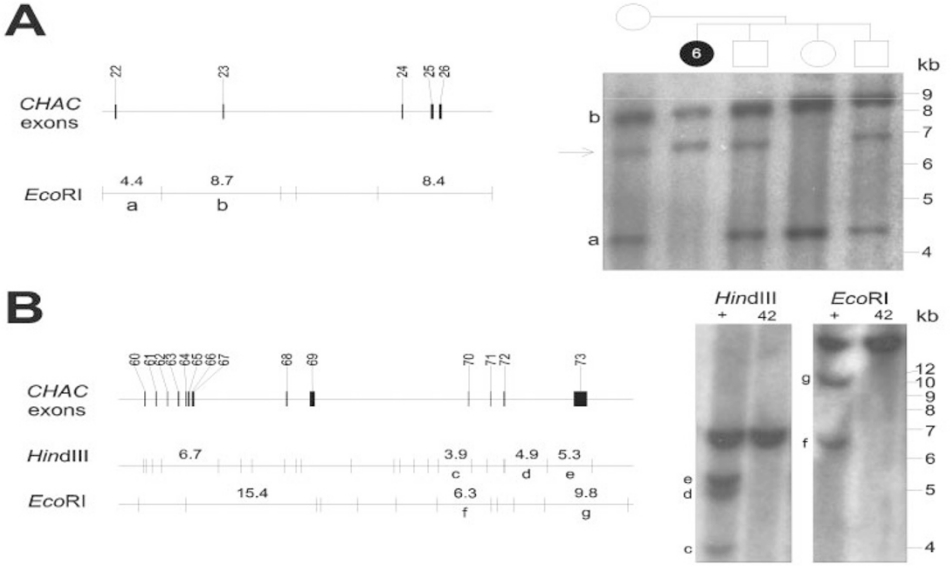

Figure 1 Southern blot analysis of gross deletions in *CHAC*. (**A**) analysis of proband 6 EX23del mutation. A 563-bp cDNA probe, containing exons 22–26, was used. A scaled restriction map of the normal locus is shown on the left; hybridisation of the probe to an *Eco*RI digest of the proband family is shown on the right. The abnormal restriction fragment is arrowed; restriction fragments that are absent in the proband are labelled a and b. (**B**) analysis of proband 42 EX70_73del mutation. A 710-bp cDNA probe, containing exons 65–68 and 70–73, was used. Scaled restriction maps of the normal locus are shown on the left; hybridisation of the probe to *Hin*dIII and *Eco*RI digests of a control (+) and proband 42 is shown on the right. Restriction fragments that are absent in the proband are labelled c–g.

Table 2 Exonic polymorphisms identified in the *CHAC* gene

Exon	*DNA change*[a]	*Protein change*[a]	*Frequency in ChAc proband chromosomes*
13	1020A>G	None[b]	1/106
18	1695C>G	F565L	2/106
26	2693T>C	V898A	1/106
26	2814A>G	none	1/106
29	3045G>A	none	12/106
38	4469G>A	R1490K	1/106
39	4760A>G	Y1587C	0/106[c]
44	5583A>G	none	15/106
45	5917G>A	V1973I	5/106
46	6027C>T	none	13/106
48	6492T>C	none	25/106
48	6531C>T	none	1/106
54	7457T>C	I2486T	1/106
63	8571T>C	none	14/106
67	9069A>G	none	50/106
73	9515C>T	P3172L	2/104

[a]Nucleotides and amino acids are numbered acording to the cDNA sequence of *CHAC* isoform a reported by Rampoldi *et al* ([11]; Genbank accession no. NM_033305), with the adenosine of the initiation codon assigned position 1. [b]Possibly affects RNA splicing: see Results. [c]Mutation found in a healthy member of family CHAC6.[10]

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 → GAG**GT**AAAT. The new sequence scores 84 on the SpliceView program as a new splice-donor site (the splice-donor site in intron 13 that is usually used scores 79). Proband 18 is heterozygous for this change and an S539X mutation in exon 18 (Table 1). 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.

Discussion

Mutation analysis of the *CHAC* gene in a cohort of 43 unrelated probands revealed 57 different mutations likely to cause chorea-acanthocytosis in 39 pedigrees. Fifty-four of these disease mutations have not previously been reported, indicating a strong allelic heterogeneity with no single mutation causing the majority of ChAc cases in the population. For four probands in this study with typical symptoms, no disease mutations in *CHAC* were detected. As there is little evidence for locus heterogeneity, it could be that the mutations are located within the 5′ or 3′ untranslated regions, or within intronic sequences not screened in this study; small deletions and rearrangements preventing amplification of one allele will also not be detected with this methodology. For the seven probands with only one heterozygous mutation detected, this is almost certain to be the case. Since three gross deletions have already been identified ([12] and present study) and the *CHAC* gene is spread over such a large genomic region (about 250 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.

CHAC mutation distribution

Figure 2 shows that there does not seem to be any 'clustering' of mutations identified in *CHAC* so far. Nine mutations in this study (Table 1) and two from our previous screening[11] are predicted to affect only transcript A, which contains exons 70–73. Proband 42 is in fact homozygous for a transcript A-specific mutation, namely the deletion of exons 70–73 (Figure 1). Given that this patient displays symptoms typical for ChAc, we can deduce that these exons are essential for some functions of chorein and that transcript B, which lacks these exons, cannot compensate for it.

Recurrent *CHAC* mutations

Three mutations reported here were found in multiple probands within this study, or had been identified in another family in a previous study in this laboratory. Probands 12 and 39 are heterozygous and proband 27 is homozygous for the mutation R208X, which was also found in family CHAC3 reported earlier.[11] Probands 1 and 34 have mutations (R3037X and 9429_9432del, respectively) that had initially been identified in families CHAC7 and CHAC11.[11] As we could not perform haplotype analysis on certain probands in this study, it is not possible to determine whether they are related. Family CHAC3 and proband 12 both originate from Italy and so it is conceivable that their mutation was inherited from a common ancestor. However, proband 27 originates from Japan and proband 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 R208X (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. The same explanation is appropriate for the probable recurrence of the R3037X (9109C>T) mutation in the Mexican family CHAC7 and the German proband 1; and may also explain the preponderance of R>X substitutions in this cohort. Proband 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 smal-

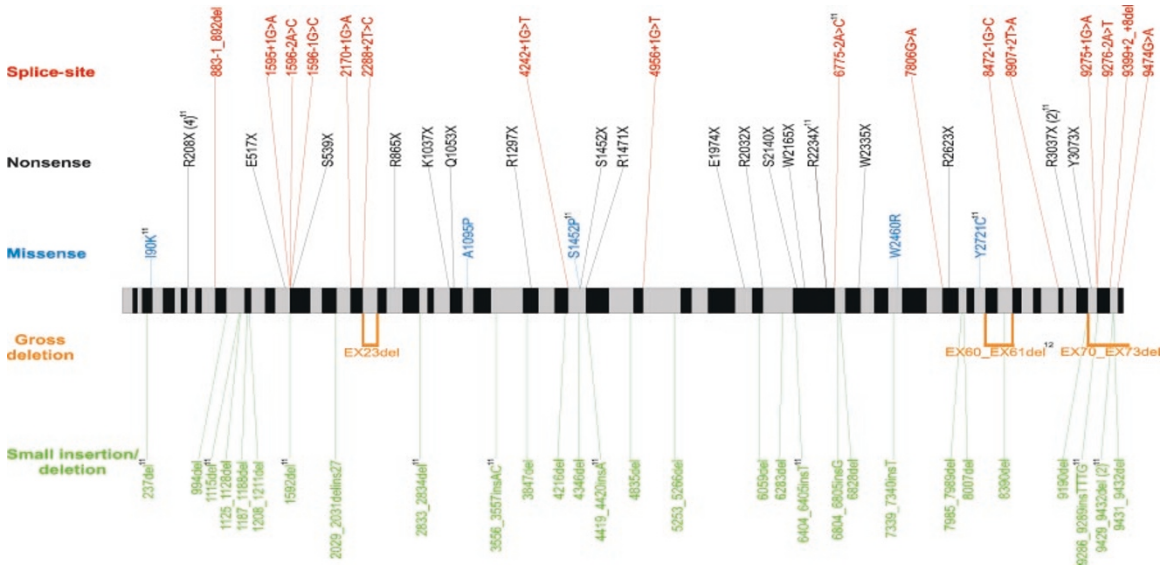

Figure 2 Spectrum of *CHAC* mutations identified. The coding sequence of *CHAC* transcript A is illustrated; exons are denoted by grey and black bars. Mutations identified in previous studies are referenced. The number of probands in which identical mutations were identified is shown in parentheses.

ler deletion within the same region in proband 29 (9431_9432del, deletion of the dinucleotide AG) supports the latter hypothesis.

Predicted effects of *CHAC* mutations on function

Forty of the 57 *CHAC* mutations identified in this study (70%) are nonsense or insertion/deletion mutations that lead to PTCs and therefore predict absence or marked reduction of mutated *CHAC* transcript via nonsense-mediated decay. This is in accordance with the recessive inheritance of chorea-acanthocytosis, where lack of chorein is the primary cause of the disease.

Fifteen out of the 57 *CHAC* disease mutations described here (26%) are changes affecting splicing. A likely explanation for this relatively high percentage of splice-site mutations is the large number of splice sites contained within this gene: the 143 splice sites of *CHAC* 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 two of the *CHAC* mutations that are predicted to cause disease in this study (2/57, 4%) are missense mutations. They are non-conservative substitutions and were not found in 384 control chromosomes, which suggests that they do not represent benign polymorphisms. The fact that just two out of nine amino-acid variants reported here (Tables 1 & 2) are believed to be pathogenic is perhaps an indication that much of chorein is tolerant to substitution.

Although it has been established that mutations in *CHAC* 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. There seems to be no significant genotype-phenotype correlation and so it is not possible to identify what clinical consequences are seen when different parts of the protein are ablated. As there are also no identifiable domains to give clues to chorein's normal function(s), information has to be derived from studies on its homologues. Only two known proteins show significant homology to chorein. TipC, which is believed to play a role in *Dictyostelium*'s morphogenetic signalling pathway,[17] has homology mainly to chorein's N- and C- termini. Vps13p in *Saccharomyces cerevisiae* is the closer homologue: it shows 41% similarity with chorein, rising to 54–58% in the N- and C-terminal regions. Yeast with mutations in *vps13* (*soi1*) show mislocalisation of some *trans*-Golgi network (TGN) and vacuolar proteins.[13] Further studies suggested that Vps13p functions at two steps in the cycling of Kex2p and other proteins between the TGN and the prevacuolar compartment.[18]

If we extrapolate from these findings we can speculate that chorein may play a similar role in protein trafficking. Proteins which could be sorted by chorein include furin, the mammalian homologue of Kex2p,[19] and cytoskeletal components. It is known that mice lacking the erythrocyte membrane protein 4.1 exhibit neurobehavioural deficits and abnormal erythrocyte morphology.[20]

Perhaps mislocalisation of a similar protein or proteins in the absence of functional chorein could cause the haematological and neurological abnormalities seen in ChAc.

With only data from *CHAC* homologues to draw on, it is clear that determination of chorein function is in its infancy. In this study we identified 54 novel mutations in the *CHAC* gene. 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.

Acknowledgements

The authors thank the patients and their families for their participation; and C Fletcher and L Lonie for DHPLC analysis. This work was supported by the Wellcome Trust. C Dobson-Stone is supported by a Wellcome Trust Prize Studentship. L Rampoldi was supported by a postdoctoral fellowship of the Italian Telethon and by a fellowship of the University of Padua. M Brin has received support from the Bachmann-Strauss Foundation. AP Monaco is a Wellcome Trust Principal Research Fellow.

References

1 Levine IM, Estes JW, Looney JM: Hereditary neurological disease with acanthocytosis. A new syndrome. *Arch Neurol* 1968; **19**: 403–409.
2 Critchley EM, Clark DB, Wikler A: An adult form of acanthocytosis. *Trans Am Neurol Assoc* 1967; **92**: 132–137.
3 Hardie RJ, Pullon HW, Harding AE *et al*: Neuroacanthocytosis. A clinical, haematological and pathological study of 19 cases. *Brain* 1991; **114** (Pt 1A): 13–49.
4 Kuroiwa Y, Ohnishi A, Sato Y, Kanazawa I: Chorea acanthocytosis: clinical pathological and biochemical aspects. *Int J Neurol* 1984; **18**: 64–74.
5 Metzer WS: Neuroacanthocytosis with autosomal dominant inheritance, normal serum CK and preserved myotatic reflexes. *Mov Disord* 1990; **5** (Suppl 1): 92, (Abstract).
6 Bonaventura I, Matias-Guiu J, Cervera C, Codina Puiggros A: Neuroacanthocytosis syndrome, apraxia of eyelid opening, and progressive supranuclear palsy. *Neurology* 1986; **36**: 1276.
7 Peppard RF, Lu CS, Chu NS, Teal P, Martin WR, Calne DB: Parkinsonism with neuroacanthocytosis. *Can J Neurol Sci* 1990; **17**: 298–301.
8 Kartsounis LD, Hardie RJ: The pattern of cognitive impairments in neuroacanthocytosis. A frontosubcortical dementia. *Arch Neurol* 1996; **53**: 77–80.
9 Feinberg TE, Cianci CD, Morrow JS *et al*: Diagnostic tests for choreoacanthocytosis. *Neurology* 1991; **41**: 1000–1006.
10 Rubio JP, Danek A, Stone C *et al*: Chorea-acanthocytosis: genetic linkage to chromosome 9q21. *Am J Hum Genet* 1997; **61**: 899–908.
11 Rampoldi L, Dobson-Stone C, Rubio JP *et al*: A conserved sorting-associated protein is mutant in chorea-acanthocytosis. *Nat Genet* 2001; **28**: 119–120.
12 Ueno S, Maruki Y, Nakamura M *et al*: The gene encoding a newly discovered protein, chorein, is mutated in chorea-acanthocytosis. *Nat Genet* 2001; **28**: 121–122.
13 Redding K, Brickner JH, Marschall LG, Nichols JW, Fuller RS: Allele-specific suppression of a defective trans-Golgi network (TGN) localization signal in Kex2p identifies three genes involved in localization of TGN transmembrane proteins. *Mol Cell Biol* 1996; **16**: 6208–6217.
14 Johnson SE, Dahl A, Sjaastad O: Progressive pseudobulbar paresis, early choreiform movements, and later rigidity: appearance in two sets of dizygotic twins in the same family. *Mov Disord* 1998; **13**: 556–562.
15 Dobson-Stone C, Cox RD, Lonie L *et al*: Comparison of fluorescent single-strand conformation polymorphism analysis and denaturing high-performance liquid chromatography for detection of EXT1 and EXT2 mutations in hereditary multiple exostoses. *Eur J Hum Genet* 2000; **8**: 24–32.
16 Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning. A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press, 1989.
17 Stege JT, Laub MT, Loomis WF: Dip genes act in parallel pathways of early ictyostelium development. *Dev Genet* 1999; **25**: 64–77.
18 Brickner JH, Fuller RS: SOI1 encodes a novel, conserved protein that promotes TGN-endosomal cycling of Kex2p and other membrane proteins by modulating the function of two TGN localization signals. *J Cell Biol* 1997; **139**: 23–36.
19 Nakayama K: Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. *Biochem J* 1997; **327** (Pt 3): 625–635.
20 Walensky LD, Shi ZT, Blackshaw S *et al*: Neurobehavioral deficits in mice lacking the erythrocyte membrane cytoskeletal protein 4.1. *Curr Biol* 1998; **8**: 1269–1272.
21 den Dunnen JT, Antonarakis SE: Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat* 2000; **15**: 7–12.

Appendix 1 *CHAC* primers[a]

Exon	*Sequence of primer pair*[b]	*Amplicon size* (bp)	*$MgCl_2$ Concentration/AT* (mM/°C)	*Application*
1	**F: ACCGCCTCCGTCTCTCGCTGG**	362	2/62	DHPLC
	R: GAAGCCCCGAGAACGCAG			
11	F: CCACACATGGTTCATCTGACTG	343	3/54	DHPLC
	R: TTGCTTTTATCAAACAATAATTCAC			
12	**F: CCCTAAAAAGTCAGTAATGTAAC**	360	2.5/58	DHPLC
	R: CTCAATAAGCTAAAGCAGCTAC			
16	**F: TTTATTAAAGGCAGATACATTGTG**	207	3/58	Sequencing
	R: TATATTCTAAGGTATAAGGCAAG			
18	F: ATTGTATGTCTTCTTTGATGGG	434	2.5/58	DHPLC
	R: CTATACTATTCACTGTCCTCTG			
18	**F: GTCTTCATTGGATGATGCAATGCCAC**	188	2.5/58	*Mwo*I RFLP
	R: CATCATTAACACAAACTCAAAGCAAAC			
19	**F: AATTTAGTTTGCTTTGAGTTTGTG**	367	2.5/54	DHPLC
	R: CTAGGTATGAACTTTCTGAGTG			
21	F: TACTTGTGAATTTATCCTCTCC	292	2.5/58	DHPLC
	R: ATCACTTCAACAAAATCAATCAAC			
24	**F: GCGGTGAATAAACATTTTTGTGC**	230	2.5/62	DHPLC
	R: CTTCCTCTGTGTGGTCCTTTGAG			
27	**F: CATTTATTTTCAGTTACCTTTCC**	254	2.5/54	DHPLC
	R: TACAAAGAGATTATGTTACAGTG			
28	F: TGATTGTATATGTATGTCTTCC	415	2.5/58	DHPLC
	R: GCTGATTTTTCCTCTGATTGC			
29	**F: TCTAACTGGAATTGTAAGATCC**	444	2.5/54	Sequencing
	R: CCTGGCATATAATGAATCCTTC			
31	**F: AAATTTCACAAAAAGCAAATTAAAAG**	148	2.5/50	*Bst*XI RFLP
	R: GAATCTAAAACAATTATATTCCATAGC			
38	**F: AAAAAGACATGATGGATATAAAGTGC**	166	2.5/58	*Bsg*I RFLP
	R: TGCACACTGGTTTCTACAGCAG			
54	**F: TGCCTATAGATTTGGGGGAAAAGTC**	135	2.5/58	*Hinc*II RFLP
	R: GCTCTGCTAACTCTGCTTTCTCAC			
69	F: TAATTACAGTCTGAGTCATAGC	317	2.5/54	DHPLC
	R: TATGGGCATTATGGTATAAGTC			

[a]Primers that were specifically designed for this study. All other primers are reported in our previous study.[11] [b]Primers that were redesigned for this study are highlighted in bold. AT, annealing temperature; DHPLC, denaturing high-performance liquid chromatography; F, forward; R, reverse; RFLP, restriction fragment length polymorphism analysis.