Linkage and sequence analysis indicates that *CCBE1* is mutated in recessively inherited Generalised Lymphatic Dysplasia.

Fiona Connell[1 *] Kamini Kalidas[1*], Pia Ostergaard[1*], Glen Brice[2], Tessa Homfray[2], Lesley Roberts[3], David J Bunyan[4], Sally Mitton[5], Sahar Mansour[2], Peter Mortimer[6], Steve Jeffery[1]
Lymphoedema Consortium [7].

Corresponding author:

1. Professor Steve Jeffery. Medical Genetics Unit. Clinical Developmental Sciences. St George's University of London, Cranmer Terrace, London, United Kingdom. SW17 0RE

Co-authors:

1. Medical Genetics Unit. Clinical Developmental Sciences. St George's University of London, Cranmer Terrace, London, United Kingdom. SW17 0RE

2. South West Thames Regional Genetics Unit, St George's University of London, Cranmer Terrace, London, United Kingdom. SW17 0RE.

3. Harris Birthright Unit, Obstetrics and Gynaecology, King's College Hospital, Denmark Hill, London, United Kingdom, SE5 9RS

4. Wessex Regional Genetics Laboratory, Salisbury District Hospital, Salisbury, Wiltshire, United Kingdom, SP2 8BJ.

5. Child Health, Clinical Developmental Sciences. St George's University of London, Cranmer Terrace, London, United Kingdom. SW17 0RE

6. Cardiac and Vascular Sciences (Dermatology), St George's University of London, Cranmer Terrace, London, United Kingdom. SW17 0RE.

7. Authors are listed in the Acknowledgements

* These 3 authors made equal contributions to this work, and are listed in alphabetical order.

Abstract
Generalised lymphatic dysplasia (GLD) is chararacterised by extensive peripheral lymphoedema with visceral involvement. In some cases it presents in utero with hydrops fetalis. Autosomal dominant and recessive inheritance has been reported. A large, non-consanguineous family with three affected siblings with generalised lymphatic dysplasia is presented. One child died aged 5 months, one spontaneously miscarried at 17 weeks gestation, and the third has survived with extensive lymphoedema. All three presented with hydrops fetalis. There are seven other siblings who are clinically unaffected.

Linkage analysis produced two loci on chromosome 18, covering 22Mb and containing 150 genes, one of which is *CCBE1*. A homozygous cysteine to serine change in CCBE1 has been identified in the proband, in exon 3, in a residue that is conserved across species. High density SNP analysis revealed homozygosity (a region of 900kb) around the locus for *CCBE1* in all three affected cases. This indicates a likely ancestral mutation that is common to both parents; an example of a homozygous mutation representing Identity by Descent (IBD) in this pedigree. Recent studies in zebrafish have shown this gene to be required for lymphangiogenesis and venous sprouting and are therefore supportive of our findings. In view of the conserved nature of the cysteine, the nature of the amino acid change, the occurrence of a homozygous region around the locus, the segregation within the family, and the evidence from zebrafish, we propose that this mutation is causative for the generalised lymphatic dysplasia in this family, and may be of relevance in cases of non-immune hydrops fetalis.

Introduction.

Lymphoedema is chronic, and often progressive, swelling due to failure of lymph drainage of protein-rich fluid from the interstitium, in circumstances in which capillary drainage is not increased (Mortimer 1998). Lymphoedema can be primary or secondary (acquired). Primary lymphoedema is a chronic oedema caused by a developmental abnormality of the lymphatic system (Mortimer 1995). It usually affects the extremities as a result of abnormal regional lymphatic failure, although visceral lymphatic drainage can also be impaired. The accumulation of fluid in the interstitial spaces occurs due to an anatomical or functional defect in the lymphatic vessels. Impairment of the lymphatic drainage system can occur as a non-syndromic mendelian condition or as part of a more complex syndromic disorder (Ferrell and Finegold 2008).

Lymphoedema can be debilitating and disfiguring. It causes physical and psychological morbidity, and in some cases can be life-threatening. The burden of this disease and its complications is considerable. Treatment options are limited and only palliative.

In recent years there has been considerable progress made in the understanding of the molecular pathways underlying lymphangiogenesis, with mouse knockouts

uncovering a range of gene products, from the master switch, PROX1 (Wigle and Oliver 1999), through to the guidance molecules Ephrin B4 and ephrin B2 (Makinen et al 2005). In humans, only three causative genes have been identified for disorders where lymphoedema is the primary phenotype; *VEGFR3* in Milroy disease (Ferrell et al 1998, Karkkainen et al 2000, Irrthum et al 2000), *FOXC2* for Lymphoedema distichiasis (Fang et al 2000, Bell et al 2001, Erickson et al 2001, Brice et al 2002)) and mutations in *SOX18* are responsible for causing the rare syndrome, hypotrichosis-lymphoedema-telangiectasia (Irrthum et al 2003). There are many other forms of primary lymphoedema where the genetic cause is unknown and the phenotype not well delineated. Widespread peripheral lymphoedema associated with systemic involvement is termed generalised lymphatic dysplasia (GLD). It can present *in utero* with ascites, pleural/pericardial effusions and hydrops fetalis. In 1989, Hennekam et al described an autosomal recessive condition in which members of a consanguineous family were reported with congenital severe peripheral lymphoedema, intestinal lymphangiectasia, facial anomalies, seizures, mild growth retardation and mental retardation (Hennekam et al. 1989). The phenotype of Hennekam syndrome has since been expanded and the degree of mental retardation is reported to be very variable, even within families, ranging from near normal development to severe delay (Forzano et al. 2002;Van Balkom et al. 2002;Yasunaga et al. 1993). The generalised lymphatic maldevelopment that characterises this syndrome predominantly affects the limbs and bowel, but can also manifest in the face, genitalia and other organs (pleura, pericardium, thyroid and kidneys) (Forzano et al. 2002;Van Balkom et al. 2002). The lymphoedema is usually congenital, can be asymmetrical and, often, gradually progressive (Van Balkom et al. 2002). Additional features include, congenital heart defects, vascular anomalies, craniosynostosis, congenital glaucoma, polysplenia, ear anomalies, hearing loss, primary hypothyroidism, pyloric stenosis, camptodactyly, rectal prolapse and renal malformations (Ali-Gazali et al. 2003;Angle and Hersh 1997;Cormier-Daire et al. 1995;Scarcella et al. 2000;Van Balkom et al. 2002). A genetic cause of Hennekam syndrome has yet to be identified, and there may well be genetic heterogeneity for this syndrome, but it is families such as the one described in this report (see below) that may lead to further insight into this syndrome and clarification of the spectrum of the phenotype of this rare condition.

A family with non consanguineous parents, seven unaffected children and three offspring affected with a generalised lymphatic dysplasia has been ascertained and genetically investigated. The proband has extensive lymphoedema of all parts of the body, and DNA was available from the two other affected children, both deceased. Linkage analysis identified two loci on chromosome 18, with 150 genes within them. Sequence analysis of these genes was undertaken.

Family profile: The family was ascertained via the joint lymphoedema/genetic clinic at St George's Hospital, London, UK. The proband, II3**,** (Fig 1) is the child of non-consanguineous parents. Antenatal hydrops, with pleural effusions and ascites, was diagnosed at16 weeks gestation. The increasing ascites and polyhydramnious required drainage via a peritoneal shunt at 33 weeks gestation.The antenatal karyotype result was 46 XX (confirmed postnatally) and fetal echocardiogram was normal.

The proband was born at 35 weeks gestation weighing 3.8kg. At birth, she was oedematous and was admitted to the neonatal unit requiring ventilation and inotropic

support. Her albumin following birth was 18g/L (normal range 35-48g/L) and she received regular albumin infusions. Liver function was otherwise normal. She had an enlarged clitoral glans but urinary sex hormone profiles, a short synacthen test, and metabolic profile were all normal. After commencing medium chain triglyceride (MCT) enteral nutrition at the age of 32 days, she developed severe diarrhoea. Bowel biopsies taken at seven weeks of age did not reveal evidence of intestinal lymphangiectasia but subsequent bowel histology showed evidence of lymphatic dilatation and inflammation. She continued on a treatment regime of albumin infusions and total parenteral nutrition. At this time she had severe peripheral oedema affecting the legs, feet, arms, hands and face (including conjunctiva). At the age of three months she developed a severe stridor. Laryngotracheobronchoscopy showed a narrowed trachea, compression of the trachea by oedema and pulmonary hypoplasia. Respiratory difficulties secondary to upper respiratory tract obstruction and reduced lung volume were ongoing in the first year of life. Age six years, she continues to have widespread, generalised lymphoedema, affecting the whole body (see figure 2). Recurrent ascites is an ongoing problem, for which varying treatments have been tried, including albumin infusions, diuretics (spironolactone) and octreotide infusions. Her serum albumin is persistently low (approximately 19g/L). She is kept on an MCT diet as treatment for intestinal lymphangiectasia. Manual lymphatic drainage and multi-layer bandaging techniques have been used to try and reduce the peripheral oedema. She has dysmorphic facies consistent with past *in-utero* oedema (epicanthic folds, depressed nasal bridge) and ongoing facial lymphoedema. Her development has been essentially normal, allowing for some delay in achieving early developmental milestones given the severity of her ongoing illness. She continues on thyroxine treatment for hypothyroidism

Male infant (II.1 Fig 1) was born at 29 weeks gestation with a birth weight of 2.78kg. Hydrops was observed on the 20 week gestation ultrasound scan. The pregnancy was further complicated by polyhydramnios. An antenatal echocardiogram was normal but postnatally he developed hypertrophic cardiomyopathy. He was oedematous at birth (normal serum albumin), with chylous ascites and chylous pleural effusions, and had severe pulmonary hypoplasia with pulmonary lymphangiectasia. Oral feeding could not be established. He died at the age of five months having never been discharged from the neonatal unit. No post mortem was carried out. The differential diagnoses in this infant included Noonan syndrome as hydrops and hypertrophic cardiomyopathy are associated with this diagnosis. However, the subsequent family history and phenotype in the other two affected family members made this diagnosis unlikely.

Male infant (II.2 Fig 1)): an ultrasound scan of this fetus at 16/40 weeks gestation revealed hydrops fetalis. The pregnancy spontaneously miscarried at 17 weeks gestation. A post mortem examination showed severe non-immune hydrops, with normal heart and lungs, and no other structural abnormalities reported. There were 4mls of blood stained ascitic fluid in the peritoneal cavity, and 2ml straw coloured fluid in the pleural cavities.. Infective causes of hydrops were ruled out, and lysosomal enzymes were normal, making storage diseases unlikely.
The occurrence of three pregnancies (two male and one female) affected by hydrops, wirh no other family history, is suggestive of autosomal recessive inheritance of a generalised lymphatic dysplasia (GLD).

Methods.

SNP chip array and linkage analysis. DNA was extracted from peripheral blood using a standard chloroform ethanol procedure. There was no blood available from the fetus or baby that died, and DNA was taken from tissue samples in these cases. The linkage SNP microarray analysis was performed at the Institute of Child Health, London, using standard procedures, with an Affymetrix 10K SNP-chip (Xba142 with 10204 SNPs). Processing of arrays was performed according to the manufacturer's protocol and call rates were all above 99%. All genotypes were checked for Mendelian inconsistencies in MERLIN. A parametric linkage test was run in MERLIN (Abecasis et al 2002) using an autosomal recessive model. Microsatellite and polyacrylamide gel electrophoresis were used to confirm and refine the two regions. This work resulted in a final list of 150 genes which were prioritised according to their functional relevance. DNA from the three affected individuals was hybridised on a 250k Affymetrix SNP chip (Part number 900767, 262,000 SNPs) at Gene Service, Nottingham, following the manufacturer's protocol. Homozygous regions on chromosome 18 shared by the 3 affected siblings were investigated using Exclude AR (Woods et al 2004).

Gene sequencing. Primers for candidate genes were designed using the Primer3 software (Rozen and Skaletsky, 2000). *CCBE1* has 11 exons, and primer sequences and PCR programmes are available on request. The cleaned PCR products were sequenced using BigDye Terminator v3.1 and an ABI 3130xl Genetic Analyzer. The sequencing traces were visually inspected in Chromas Lite (Technelysium Pty Ltd) and finally compared to wild type sequence using CLC Sequence Viewer 6.2 (CLC bio A/S). As no restriction site was found for the SNP in exon 3, two hundred controls (caucasian) were sequenced. Primers for exon 3 which produced allele dropout were: 3F- AATGTTTCCTGGGCACAAGT 3R-AAGACCTATATTCCATGAACATCTGA while those where amplification was successful were:
3aF ACTTCACCCCTACTTTGCTTT 3cR AGGGAGGAGGGTTGGTTCT

Mutation database analysis. The change found in CCBE1 was examined for potential effect on the protein using SIFT (sift.jcvi.org) and PolyPhen (genetics.bwh.harvard.edu/pph/)

RNA Analysis. Blood was taken from the mother using a PAXgene Blood RNA tube (PreAnalytiX). RNA was purified following the manufacturer's protocol, the quality was checked on Nanodrop. cDNA was obtained using SuperScript II Reverse transcriptase (Invitrogen). A skin biopsy was also taken from the mother, stored in RNALater (Ambion) for 48 hours in a refrigerator, and then RNA was extracted.using the same protocol as for the blood sample. Primers were designed to check for the presence of an RNA transcript from *CCBE1* in the two tissues using Primer3 software and are available upon request. Control primers for *GAPDH* were used to check the cDNA.

Results.

Using random markers in a genome wide analysis, the family was only large enough to generate a LOD score of 2.4, so could not reach the conventionally significant LOD score of 3. The severity of the phenotype meant that the chance of the clinically unaffected siblings carrying the affected genotype was remote, although it had to be considered. Linkage analysis was carried out on the data from a 10kb Affymetrix SNP chip, using MERLIN. This produced two loci on chromosome 18; one at 18p and one at 18q (Fig 3). These both gave LOD scores of 2.4. The remainder of the genome gave no loci returning a positive LOD score.

The spacing and informativity of the SNPs meant that microsatellite mapping was needed to refine the two intervals (see figure 1 and Table 1). The two regions shared by the three affected siblings were large, (approximately 17Mb for the interval on the p arm and 15Mb for that on the q arm), but shared haplotypes in II.9 and II.10 and affected siblings reduced the 18p interval to 7Mb. The region on 18q could not be further reduced.

Database mining of the intervals using NIBSC (ncbi.nlm.nih.gov/sites/entrez), Ensembl (ensemble.org) and UCSC (genome.ucsc.edu) produced the current knowledge for all genes in the intervals. Those candidates considered 'Good' or 'Average' on the basis of their biochemistry, function and distribution, are shown in Table 2. Sequencing of all the 'Good' candidates yielded no mutations, but in the 'Average ' group was *CCBE1*, which had a suggested role in calcium transport, or an EGF binding function, based on homology to other proteins. Sequence analysis identified an apparently homozygous mutation in the proband and her two affected siblings, changing a cysteine to a serine (Fig 4). The change is p.Cys75Ser (c.223T>A). This residue is conserved across species, being present in 37 of the 40 species represented in the Ensembl homologue database (See Table 3). In the three in which it was not present; stickleback, medaka and platypus, there was no equivalent residue. The PolyPhen database indicates this as a probably damaging mutation, and SIFT predicts that the change will affect protein function. Recently, ccbe1 was reported to have an 'indispensable role in lymphangiogenesis' by Hogan et al (2009) in the zebrafish. The approximate position of *CCBE1* in the 18q linkage interval is shown in Figure 1.

Familial DNA analysis using the first primer set indicated heterozygosity in the paternal haplotype but both maternal alleles were apparently T. However it was determined that the penultimate base of the reverse primer contained a SNP (found in intron 3) resulting in non-amplification of one of the mother's alleles (see Methods for primer details). Primers were therefore redesigned to span the mutation and the SNP (see Methods), and results using the second primer set showed that there was indeed both heterozygosity for the SNP and for the mutation in the mother (mutation shown in Fig 5). All members of the family were sequenced with the new primer set, and this produced the result that was expected from the haplotype data; both parents were heterozygous for p.Cys75Ser, and all three affected siblings were homozygous. Individuals II.4, II.7, II.9 and II.10 were heterozygous, and II.5, II.6 and II.8 were homozygous wildtype. This change was not found in 400 control chromosomes.

The occurrence of the same mutation in unrelated parents prompted investigation into the possibility that this homozygosity might be more extensive than simply this one base, and analysis of chromosome 18 in the three affected individuals, with a 250k genomic SNP chip, produced an area of homozygosity of about 900kb. A heterozygous SNP at rs4349256 (54,834,349 base position) delineated the upper boundary of homozygosity.

There was another heterozygous SNP at rs4940912 (55,732,132). This region of shared homozygosity contained *CCBE1* (55,250,000 – 55,513,000).

As very little is known about the expression of CCBE1, we investigated its transcription in lymphocytes and skin from the mother of the proband. If the transcript was produced it would also allow confirmation of the mutant residue in the RNA . cDNA was successfully produced from both white cells and skin, and GAPDH amplified in both. However, no product at all was obtained for CCBE1 from lymphocytes, but a strong signal was given by the skin RNA (Fig 6). Sequencing of this product showed that both alleles were expressed (data not shown).

Discussion

A homozygous mutation in the *CCBE1* gene has been identified in three siblings affected by generalised lymphatic dysplasia (GLD). The phenotype is severe, causing significant morbidity and mortality in this family. Examination of chromosome 18q in the affected individuals in the pedigree has shown that they all share a homozygous region around *CCBE1* of just under a megabase. This is strongly suggestive of a previously unknown distant relationship between the parents, and Identity by Descent (IBD) of the mutation in this family.

Ccbe1 has very recently been shown to be essential for the development of the lymphatics in zebrafish (Hogan et al 2009). Prior to this little was known about the function of this gene. Indeed, its name; Collagen and Calcium Binding EGF domain 1, derives from the fact that there are elements within the protein that show resemblance to other known protein family domains, rather than any knowledge of its particular function (see Figure 7). The zebrafish mutant full of fluid (fof), which was produced by mutagenesis, was found to have a homozygous mutation in exon 4 of the ccbe1 gene (Hogan et al 2009). This led to an absence of the thoracic duct and longitudinal lymphatic vessels, and such severe oedema that only three of 28 embryos survived to 36 days post fertilization. As ccbe1 was not expressed in lymphatic tissue, but seemed to be found in the routes where lymphatics were developing, it was proposed that it might act as part of the guidance system for budding lymphatics in rather the same way that ephrin B2 and ephrin B4 have been suggested to be involved in such a process in mice (Makinen et al 2005).

The mutation in the zebrafish causes a change of aspartic acid to glutamic acid at residue 162. This is a change in size of R group but is otherwise non-conservative. The amino acid is conserved across species, however, with the exception of the cat where there is no equivalent residue. The base change in our report is in a more N terminal part of the molecule; the function of which is currently unknown. The exon structure of the gene, the cysteine residues, and the two proposed domains that show similarity to known protein superfamilies are shown in Fig 7. The suggested homology for a stretch of CBBE1 from amino acid 71 to 219 with the growth factor domain superfamily is indicated in Fig 7, and contains the cysteine that is mutated in the patients reported here. This cysteine rich region is found in a number of proteins that have an involvement in signal transduction by receptor tyrosine kinases. Insulin growth factor binding protein (IGFBP) is a member of this family, and controls the distribution, function and activity of insulin-like growth factors I and II, which are key regulators of cell proliferation. The highest conservation of IGFBPs is in the N-terminal Cys-rich IGF-binding domain, where there are 10-12 conserved cysteine residues. If CCBE1 is involved with guidance and lymphatic proliferation, it is quite likely that production of local growth factors would be essential

for this process. The cysteine to serine change seen in this family is significant, as cysteine residues are often involved in inter or intramolecular disulfide bond formation. This residue is highly conserved, and is only absent in three of 41 species, where there is not an equivalent amino acid in the orthologue.

We wanted to look at two reasonably readily available sources of protein and RNA, namely blood and skin. Nagase et al,(2001) who cloned the cDNA, showed tissue expression in adult lung, liver and kidney, while Hogan et al (2009) showed that there was no expression in adults in epithelial cells or blood. Our cDNA data confirms that there is no expression in lymphocytes, but shows a transcript from skin RNA. This could be a potentially useful finding for examination of mutant transcript and protein.

Future analysis of other patients with a similar widespread generalised lymphatic dysplasia, and with a family history of hydrops fetalis, for mutations in *CCBE1* is clearly a priority.

In summary, a mutation of a highly conserved cysteine residue in exon 3 of *CCBE1* has been identified in three siblings with an autosomal recessive generalised lymphatic dysplasia. The homozygous nature of the mutation, the strong likelihood of IBD in the family, and the data from zebrafish, make it likely that this change is pathogenic. This is the first gene in which mutations cause extensive generalised lymphoedema in humans, rather than the peripheral disease associated with the currently known genes, *VEGFR3* and *FOXC2*. The identification of such a gene underscores the validity of human genetic approaches to identifying disease causing mutations. The phenotype is extremely severe, and can be lethal, and investigations into the function and nature of action of the gene product in humans may well have implications in understanding the nature of non-immune hydrops fetalis, and in the longer term the possibility of in utero therapy for this important and currently ill understood condition.

Acknowledgements.

KK, and FC were supported by the British Heart Foundation. PO is supported by the British Skin Foundation. We would like to thank the Reece Spence fund for their support. Much of this work was carried out in the Biomics Unit at SGUL. Members of the Lymphoedema Consortium not given in the author list are: Prof Alberto Smith, Professor Kevin Burnand (St Thomas') Dr Anne Child (SGUL), Dr Mansoor Sarfarazi (Connecticut). We would like to thank Dr John Crolla and Dr Shuwen Huang (Wessex Regional Genetic Laboratory) for help in analyzing the region with the mutation.

References.

Abecasis GR, Cherny SS, Cookson WO and Cardon LR (2002). Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97-101.

Al-Gazali LI, Hertecant J, Ahmed R, Khan NA, Padmanabhan R (2003). Further delineation of Hennekam syndrome. Clin Dysmorphol. 12:227-32

Angle B, Hersh JH (1997) Expansion of the phenotype in Hennekam syndrome: a case with new manifestations. Am. J. Med. Genet. 71:211-4

Bell R, Brice G, Child AH, Murday VA, Mansour S, Sandy CJ, Collin JR, Brady AF, Callen DF, Burnand K, Mortimer P, Jeffery S (2001) Analysis of lymphoedema-distichiasis families for FOXC2 mutations reveals small insertions and deletions throughout the gene. Hum. Genet. 108:546-51

Brice G, Mansour S, Bell R, Collin JR, Child AH, Brady AF, Sarfarazi M, Burnand KG, Jeffery S, Mortimer P, Murday VA (2002) Analysis of the phenotypic abnormalities in lymphoedema-distichiasis syndrome in 74 patients with FOXC2 mutations or linkage to 16q24. J. Med. Genet. 39:478-83

Cormier-Daire V, Lyonnet S, Lehnert A, Martin D, Salomon R, Patey N, Broyer M, Ricour C, Munnich A (1995) Craniosynostosis and kidney malformation in a case of Hennekam syndrome. Am. J. Med. Genet. 57:66-8

Erickson RP, Dagenais SL, Caulder MS, Downs CA, Herman G, Jones MC, Kerstjens-Frederikse WS, Lidral AC, McDonald M, Nelson CC, Witte M, Glover TW (2001) Clinical heterogeneity in lymphoedema-distichiasis with FOXC2 truncating mutations. J. Med. Genet. 38:761-6

Fang J, Dagenais SL, Erickson RP, Arlt MF, Glynn MW, Gorski JL, Seaver LH, Glover TW. Mutations in FOXC2 (MFH-1), a forkhead family transcription factor, are responsible for the hereditary lymphedema-distichiasis syndrome (2000) Am J Hum Genet 67:1382-1388.

Ferrell RE, Levinson KL, Esman JH, Kimak MA, Lawrence EC, Barmada MM, Finegold DN. (1998) Hereditary lymphedema: evidence for linkage and genetic heterogeneity. Hum Mol Genet. (13):2073-8.

Ferrell RE, Kimak MA, Lawrence EC, Finegold DN (2008) Candidate gene analysis in primary lymphedema. Lymphat. Res. Biol. 6:69-76

Forzano F, Faravelli F, Loy A, Di Rocco M (2002) Severe lymphedema, intestinal lymphangiectasia, seizures and mild mental retardation: further case of Hennekam syndrome with a severe phenotype. Am. J. Med. Genet. 111:68-70

Hennekam, R CM, Geerdink, RA, Hamel, BCJ, Hennekam, FAM, Kraus P, Rammeloo, JA, Tillemans, AAW (1989) Autosomal recessive intestinal lymphangiectasia and lymphedema, with facial anomalies and mental retardation. Am. J. Med. Genet. 34: 593-600, 1989.

Hogan BM, Bos FL, Bussmann J, Witte M, Chi NC, Duckers HJ, Schulte-Merker S. (2009) Ccbe1 is required for embryonic lymphangiogenesis and venous sprouting. Nat Genet. 41:396-8.

Irrthum A, Karkkainen MJ, Devriendt K, Alitalo K, Vikkula M. (2000) Congenital hereditary lymphedema caused by a mutation that inactivates VEGFR3 tyrosine kinase.Am J Hum Genet. 67:295-301.

Irrthum A, Devriendt K, Chitayat D, Matthijs G, Glade C, Steijlen PM, Fryns JP, Van Steensel MA, Vikkula M (2003) Mutations in the transcription factor gene SOX18 underlie recessive and dominant forms of hypotrichosis-lymphedema-telangiectasia. Am J Hum Genet 72:1470-8

Karkkainen MJ, Ferrell RE, Lawrence EC, Kimak MA, Levinson KL, McTigue MA, Alitalo K, Finegold DN. (2000) Missense mutations interfere with VEGFR-3 signalling in primary lymphoedema. Nat Genet 25:153-9.

Makinen T, Adams RH, Bailey J, Lu Q, Ziemiecki A, Alitalo K, Klein R, Wilkinson GA. (2005) PDZ interaction site in ephrinB2 is required for the remodeling of lymphatic vasculature. Genes Dev. 19:397-410.

Mortimer PS (1995) Managing lymphoedema. Clin. Exp. Dermatol. 20:98-106

Mortimer PS. The pathophysiology of Lymphoedema. Cancer. (1998) 83:2798-2802

Nagase, T.,Kikuno, R., Ohara, O. (2001) Prediction of the coding sequrnces of unidentified human genes. XXII. The complete sequence of 50 new cDNA clones which code for large proteins. DNA Res 8:319-327.

Rozen S and Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386

Scarcella A, De Lucia A, Pasquariello MB, Gambardella P (2000). Early death in two sisters with Hennekam syndrome. Am. J. Med. Genet. 93:181-3

Tammela T, Petrova TV, Alitalo K. (2005)Molecular lymphangiogenesis: new players. TRENDS in Cell Biol 15:434-441.

Van Balkom ID, Alders M, Allanson J, Bellini C, Frank U, De Jong G, Kolbe I, Lacombe D, Rockson S, Rowe P, Wijburg F, Hennekam RC (2002) Lymphedema-lymphangiectasia-mental retardation (Hennekam) syndrome: a review. Am. J. Med. Genet. 112:412-21

Wigle JT, Oliver G. (1999) Prox1 function is required for the development of the murine lymphatic system. Cell. 98:769-78.

Woods CG, Valente EM, Bond J, Roberts E. (2004) A new method for autozygosity mapping using single nucleotide polymorphisms (SNPs) and EXCLUDEAR. J. Med. Genet. 41:101-105

Yasunaga M, Yamanaka C, Mayumi M, Momoi T, Mikawa H (1993) Protein-losing gastroenteropathy with facial anomaly and growth retardation: a mild case of Hennekam syndrome. Am. J. Med. Genet. 45:477-80

Table 1. Positions on chromosome 18 for those microsatellites shown in Fig 1. The positions are for build 36.3 on NCBI Mapviewer.

5354464-5354522
6285310-6285346
7643117-7643157
7648329-7648358
7823742-7823782
8766721-8766770
12912172-12912219
13376891-13376980
14339485-14339528
14715190-14715232
15018742-15018803
17815982-17816023
20377611-20377663
21375489-21375539
22586686-22586727
23189150-23189197
24947915-24947947
27753495-27753553
29968410-29968458
35567716-35568038
37509128-37509243
42607872-42607911
44811302-44811350
46406251-46406296
47536603-47536644
47775260-47775308
49590719-49590761
57631492-57631535
59688868-59688913
61948843-61948893
63460255-63460320
63979026-63979070
64966988-64967039

Table 2. Genes designated as 'Good' and 'Average' candidates for analysis in the two linkage intervals, with a small summary to explain their inclusion.

Good candidates.

PTPN2. Protein tyrosine phosphatase, non-receptor type 2 A member of the transmembrane protein tyrosine phosphatase family that share a highly conserved catalytic motif. **10 exons.**

PTPRM. Protein tyrosine phosphatase, receptor type, M PTPs are known to be signalling molecules that regulate cell growth and differentiation**. 31 exons**.

ANKRD12. Ankyrin repeat domain-containing protein 12. Mouse homologue binds to p160, and may be part of a gene transactivation complex. **13 exons.**

VAPA. Vesicle-associated membrane protein-associated protein A. This is a SNARE; a compartmentally specific, cytoplasmically oriented integral membrane protein involved in the fusion of membranes and the transport of intracellular proteins. Localizes at the tight junction in polarized epithelial cells. Ubiquitous distribution. **7 exons.**

APCDD1. Adenomatosis polyposis coli down-regulated 1. A gene downregulated by APC in the beta-catenin T-cell factor signalling pathway in SW480 colon cancer cells. Ubiquitous expression. **6 exons.**

IMPA2. OMIM. Myo inositol monophosphate. Plays a crucial role in the phosphatidylinositol signalling pathway. **8 exons.**

CIDEA. OMIM. Cell death-inducing DFFA-like effector A. Activates apoptosis. Mice that lack functional Cidea have higher metabolic rates, higher lipolysis and higher core body temperatures when subjected to cold. These mice are also resistant to diet-induced obesity and diabetes. **5 exons.**

TUBB6. Tubulin, beta 6. (aka TUBB-5, HsT1601). Component of the cytoskeleton. **4 exons.**

AFG3L. AFG3 ATPase family gene 3-like 2 (yeast). This gene encodes a protein localized in mitochondria, closely related to paraplegin. The paraplegin gene is responsible for an autosomal recessive form of hereditary spastic paraplegia. **17 exons.**

ZNF519 Zinc Finger Protein 519. Possible transcription factor. **3 exons.**

DCC. Excellent candidate. Interacts with neural guidance molecules Nets and Slits**. 29 exons**.

CCD68. Coiled coil domain containing protein 68. **12 exons.**

TCF4. Transcription factor 4. Complexes with cJUN and beta catenin. Phosphorylation dependent interaction with cJUN regulates intestinal tumirogenesis. **20 exons**

TXNL1. Thioredoxins are small redox active proteins, with mRNA in all tissues. Transcriptional repressor binding to transcription factor B-Myb. **9 exons**

WDR7. WD repeat domain 7. Involved in cell cycle progression, signal transduction,apoptosis, and gene regulation. **28 exons**

ATP8B1. ATPase Class 1, Type 8B,member 1. Involved in ATP dependent phosholipid transport. Expressed in epithelial tissues,. Involved in enterohepatic bile acid circulation? **28 exons.**

NEDD4L. Ubiquitin protein ligase Nedd4-like. Transcript is at high levels in liver and kidney and at lower levels in brain, heart, lung, spleen, skeletal muscle, and testis. One domain, WW, binds strongly to ENAC. **30 exons.**

ALPK2. Alpha kinase 2. The gene product is believed to play a part in protein amino acid phosphorylation. **13 exons.**

MALT1. Mucosa-associated lymphoid tissue lymphoma translocation gene 1. A common translocation in tumours of the mucosa is found between this gene and AP12 (inhibitor of apoptosis). MALT1 operates downstream of BCL10, controls the catalytic activity of the I-kappa-B kinase complex, and regulates the signalling of JNK MAP kinase. **17 exons.**

ZNF532. Zinc finger protein 532. ZNF homology suggests transcription factor. **11 exons.**

LOC390858. Contains acyltransferase 3 domain and some transmembrane domains. **14 exons.**

SEC11L3. Homologue of a yeast protein that is an 18kDa catalytic subunit of the Signal Peptidase Complex which cleaves the signal sequence of proteins targeted to the ER. **6 exons.**

GRP. Gastrin releasing peptide. Increases plasma gastrin pancreatic polypeptide, glucagons, gastric inhibitory peptide, and insulin. **3 exons.**

CDH20. Similar to mouse Cdh7. Expression in placenta, adult brain, and fetal brain. Cadherins are calcium-dependent adhesive proteins that mediate cell-to-cell interaction, and are involved in the structural and functional organization of cells in various tissues. **11 exons**

RNF152. Ring Finger Protein 152. Involved in protein binding or control of transcription. **3 exons**

***PIGN*.** Phosphatidylinositol glycan, class N. Glycosylphosphatidylinositol (GPI)-anchored proteins comprise a well-characterized family of proteins that must acquire a GPI anchor and traffic from their site of synthesis, the ER, to the cell surface. **32 exons.**

ZCCHC2. The 18 residues CCHC zinc finger domain is found in eukaryotic proteins involved in RNA binding or single strand DNA binding. **15 exons.**
VPS4B. This is an AAA protein (ATPase associated with diverse cellular activities) and is involved in lysosomal/endosomal membrane trafficking. It is ubiquitously expressed. **11 exons.**
CDH7. Cadherin7. This protein is a calcium dependent cell-cell adhesion glycoprotein comprised of five extracellular cadherin repeats, contributing to the sorting of heterogeneous cell types and the maintenance of orderly structures. **12 exons.**
CDH19. Formerly given as CDH7, and shows considerable similarity to that gene. **12 exons.**

Average candidates.

RAB12. Suggested roles in cell growth, survival and differentiation, **1 exon**
***TWSG1*.** Dorsal-ventral patterning requires a conserved system of extracellular proteins. Tsg is one of these, and functions with chordin to antagonise BMP activity. **5 exons.**
RALBP1. Can catalyze the transport of glutathione conjugates and xenobiotics**. 11 exons.**
PPP4R1. Involved in protein phosphorylation on serine and threonine residues. **20 exons.**
RAB31. A small GTP-binding protein of the RAB family. **7 exons.**
TXNDC2. Probably plays a regulatory role in sperm development. **2 exons.**
NAPG. Mediates platelet exocytosis and controls the membrane fusion events. **12 exons.**
GNAL. Found in the olfactory epithelium,and in certain areas of the brain and appears to be coupled to the dopamine D1 receptor (DRD) **12 exons.**
CHMP1B. Chromatin modifying protein 1B Involved in degradation of surface receptor proteins and formation of endocytic multivesicular bodies **1 exon.**
SPIRE1. concerned with actin organization, and required for axis specification in embryos, **16 exons.**
CEP76. Centrosomal protein 76kDa. Restricted to the centrosome. **12 exons.**
TNFSF5IP1. Ubiquitous expression, and upregulated in hepatic carcinoma. **8 exons.**
SEHIL. Part of a nuclear pore complex, and, specifically localizes to kinetochores in mitosis. **9 exons.**
CEP192. Centrosomal protein 192kDa. No known function. **40 exons.**
C18orf19. Chromosome 18 open reading frame 19. **3 exons.**
RNMT. Part of the process, where 5-prime-terminal caps are formed on pre-mRNAs. **13 exons**
MC5R. Melanocortin 5 receptor. Targeted disruption of the mouse MC5R gene produced mice with a defect in thermoregulation due to decreased production of sebaceous lipids. **1 exon.**
MC2R. Melanocortin 2 receptor, encodes one member of the five-member G-protein associated melanocortin receptor family (see MCR5 above). **1 exon.**
MBD2. Binds methylated DNA and actively demethylates it. Represses transcription. **7 exons.**
POLI. Polymerase DNA Iota. Promotes replication through minor groove purine adducts. **5 exons. Av**
STARD6. Steroidogenic acute regulatory protein. Involved in cholesterol homeostasis. **6 exons.**
***C18orf54*. 8 exons.** Unknown Function.
***C18orf26*. 3 exons.** Unknown Function.
RAB27B. RABs are involved in vesicular fusion and transport. Expressed mainly in testis. **6 exons. A**
ST8SIA3. Transfers sialic acid at the termini of glycoconjugates. **4 exons.**
ONECUT2. Transcription factor from fetal retina EST. mRNA in liver and skin. **2 exons**.
FECH. Ferrochelatase. Fech is a mitochondrial enzyme. Mutations can give rise to light sensitive dermatitis (erythropoetic protoporhyria), **11 exons.**
LMAN1. Lectin mannose binding 1. A membrane mannose specific lectin. **13 exons.**
CCBE1. **Collagen and calcium binding EGF domains 1. Very little data. Could be involved in phosphate transport, or in calcium ion binding and its localisation in the cytoplasm. 11 exons.**
MC4R. Melanocortin 4 receptor. Found predominantly in the brain. **1 exon.**
KIAA1468. HEAT domains suggest involvement in protein-protein interactions. **30 exons**.
TNFRSF11A. It is the membrane-bound osteoclast differentiation factor receptor on osteoclast progenitors, **10 exons.**
PHLPP. Terminates AKT (a protein kinase) signalling by directly dephosphorylating and inactivating it. **17 exons.**
BCL2. B-cell CLL/lymphoma 2. This gene encodes an integral outer mitochondrial membrane protein that blocks the apoptotic death of some cells such as lymphocytes. **3 exons.**
FVT1.. Weakly expressed in normal haemopoetic tissues, more strongly in some T cell malignancies. Catalyzes the reduction of 3-ketodihydrosphingosine to dihydrosphingosine. **10 exons.**
HMSD. Histocompatibility (minor) serpin domain containing. Similar to SerpinB 6. **4 exons.**
C18orf4. Dermatan-sulfate epimerase-like protein precursor. No known function **3 exons**.

Table 3. Conservation of the mutated cysteine residue (shown shaded) across species. Of the 43 orthologues in Ensembl, this amino acid was not present only in those cases (platypus, cat and medaka) where there was no equivalent residue in the molecule.

Human	L	T	T	C	Y	R	K	K	C	C	K	G	Y	K	F	V	L
Chimpanzee	L	T	T	C	Y	R	K	K	C	C	K	G	Y	K	F	V	L
Lemur	-	-	-	C	Y	R	K	K	C	C	K	G	Y	K	F	V	L
Dog	L	T	T	C	Y	R	K	K	C	C	K	G	Y	K	F	V	L
Kangaroo rat	L	T	T	C	F	R	K	K	C	C	K	G	Y	K	F	V	L
Dolphin	-	-	-	-	F	R	R	K	C	C	K	G	Y	K	F	V	L
Pika (little hair)	-	-	-	-	F	R	K	K	C	C	E	G	Y	K	F	V	L
Sloth	-	T	T	C	F	R	K	K	C	C	K	G	Y	K	F	V	L
Cow	-	C	F	S	L	R	K	K	C	C	K	G	Y	K	F	V	L
Wallaby	L	T	T	C	F	R	K	K	C	C	K	G	Y	K	F	V	L
Chicken	-	-	-	-	F	R	K	K	C	C	K	G	Y	K	F	V	L
Anolis lizard	-	-	-	-	F	R	K	K	C	C	Q	G	Y	K	F	V	L
Xenopus (frog)	-	-	-	-	F	R	K	K	C	C	K	G	Y	K	F	V	L
Zebra fish	V	T	T	C	Y	R	K	K	C	C	E	G	Y	K	F	V	L
Stickleback	V	T	A	C	Y	R	K	K	C	C	K	G	Y	K	F	V	L
Fugu (puffer fish)	A	A	T	C	Y	R	K	K	C	C	K	G	Y	K	F	V	L

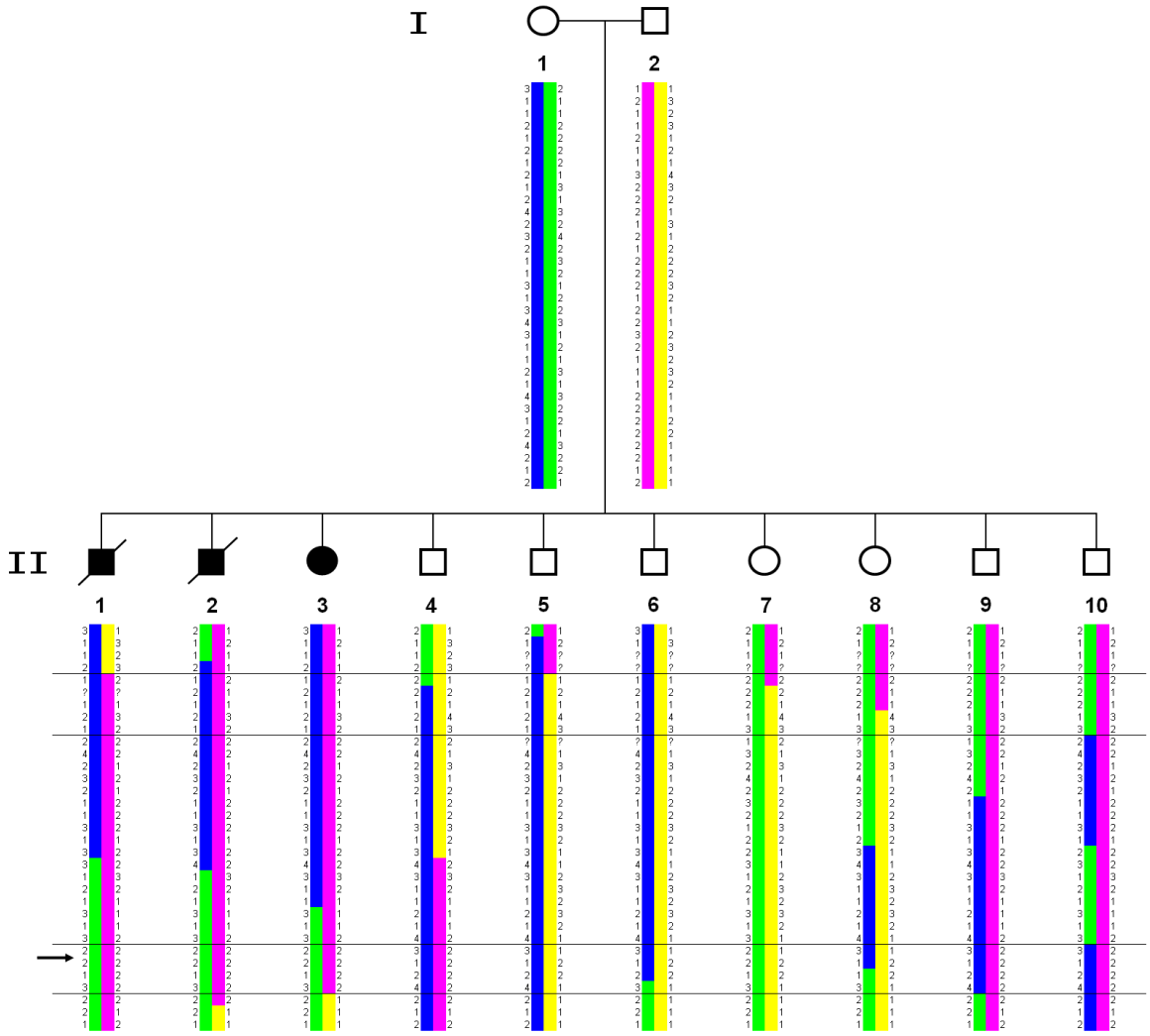


Fig 1. Haplotypes showing the delineation of the 2 loci using microsatellites. The positions of the 33 markers shown (a proportion of those used in the analysis) are given in Table 1. The two linkage intervals are given by the 2 sets of parallel lines. The position of *CCBE1* is indicated by an arrow

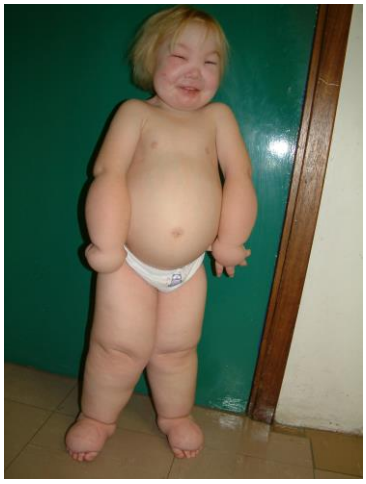


Figure 2. The proband, showing her extensive lympoedema.

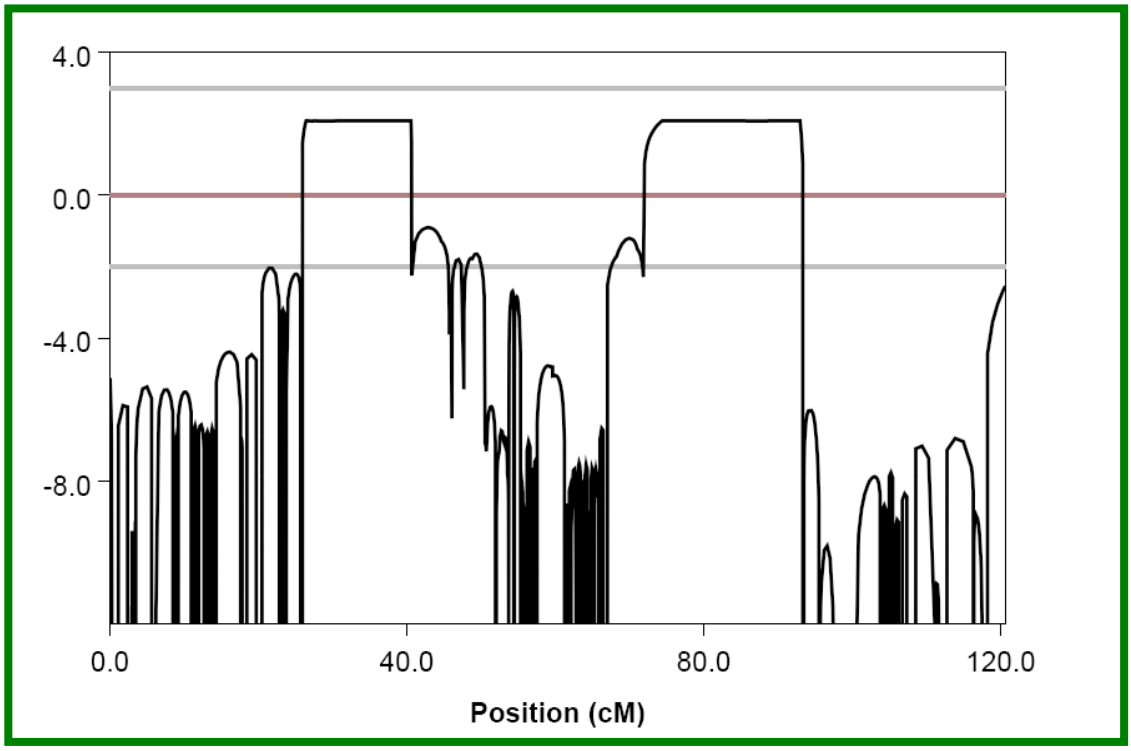
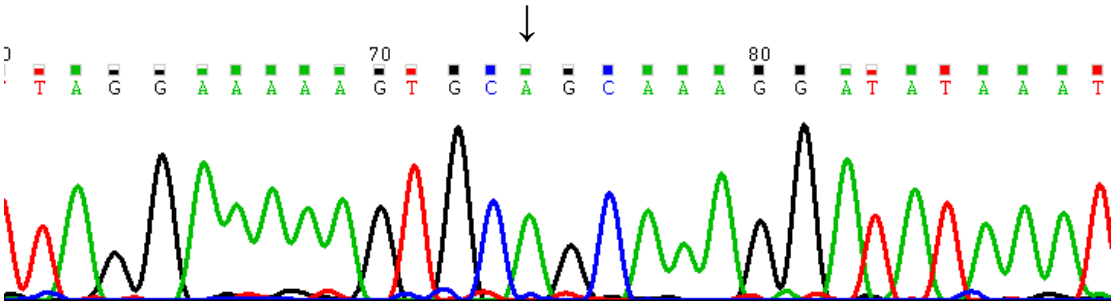
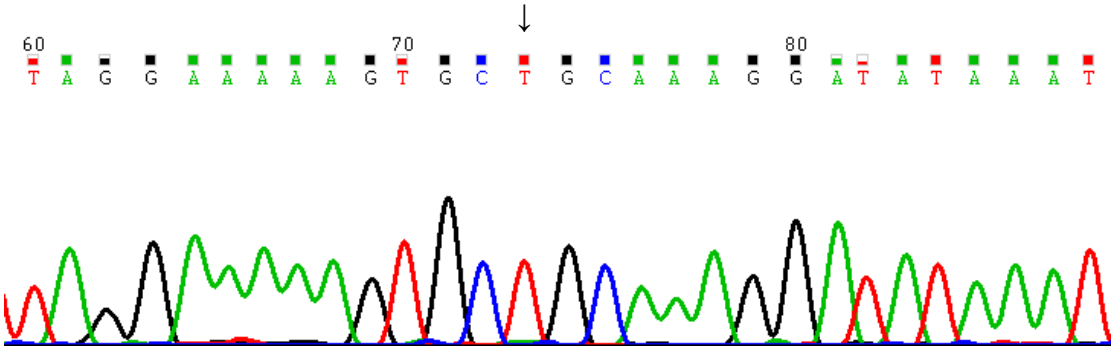


Fig 3. Lodscore results for the 2 loci on chromosome 18

Proband (II.3) *CCBE1* Exon 3



Mother (I.1) *CCBE1* Exon 3



Father (I.2) *CCBE1* Exon 3

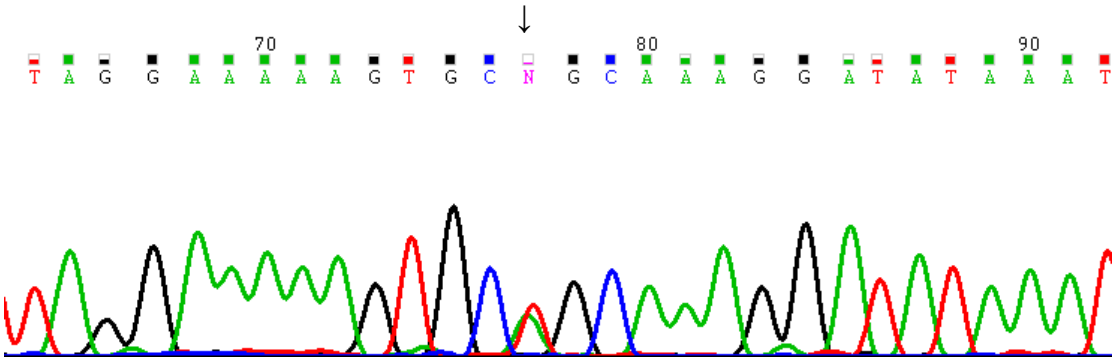


Fig 4. Sequencing traces for exon 3 showing the heterozygous allele in the father, apparent homozygosity for the wildtype T allele in the mother, and homozygous A in the affected child.

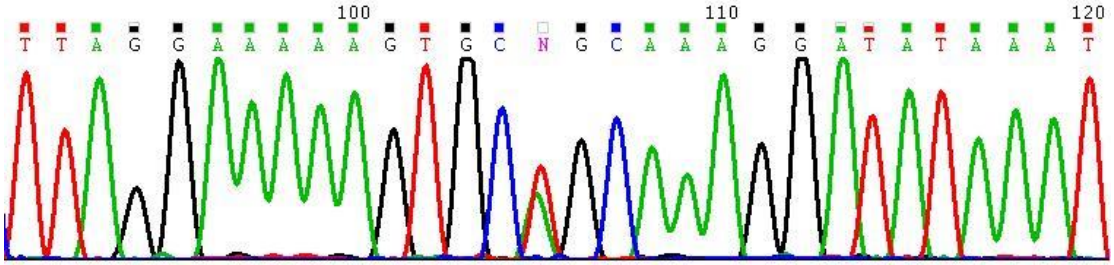


Fig 5. Sequence of exon 3 for the mother (I.1) using the second primers set, showing heterozygosity for the mutation.

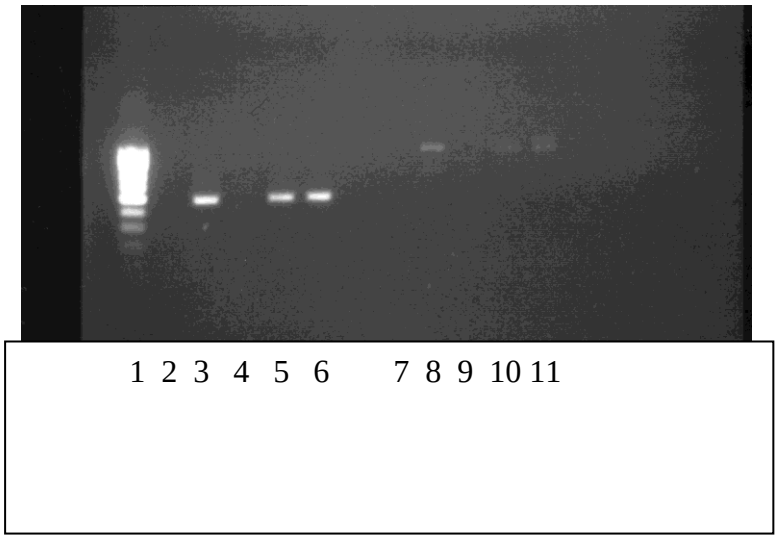


Fig 6. cDNA from mother (I.1) amplified with 2 sets of primers for CCBE1 and run on an agarose gel.

Lane 1 is a 100kb ladder Lanes 2 and 7 are negative controls. Lanes 3,4,5 and 6 are amplicons from primers in exons 1 and 6 (480bp product). 3 and 6 are from skin biopsy of I.1, 4 is from lymphocytes of I.1, 5 is a control skin biopsy.

Lanes 8-11 are amplified with primers from exons 1 and 11 (960bp product). Lanes 8 and 11 are from skin biopsy of I.1, lane 9 is from lymphocytes of I.1, lane 10 is a control skin biopsy.

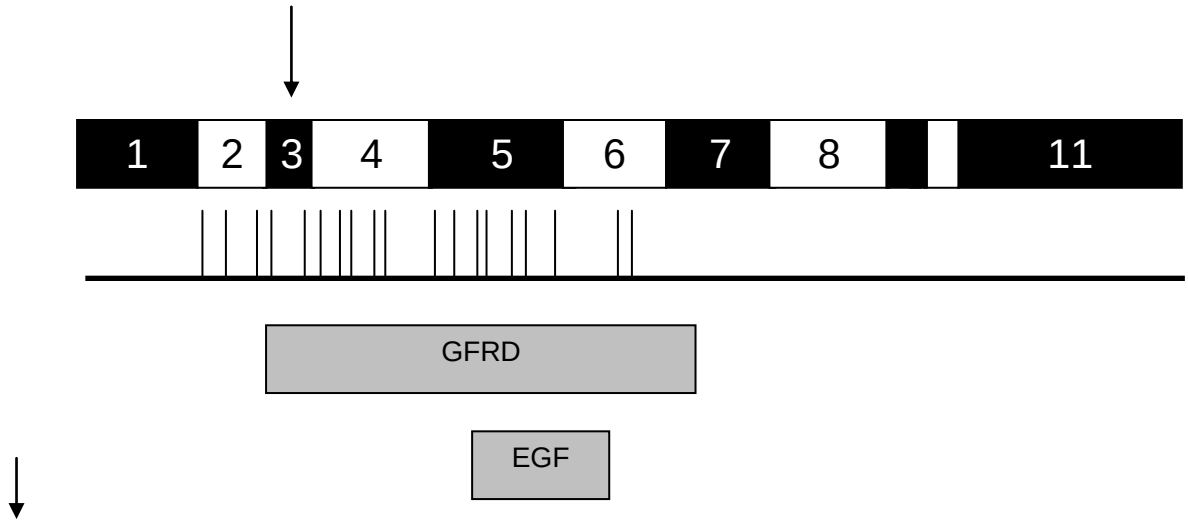


Fig7. Schematic of CCBE1, indicating the 11 exons in the top line, the position of the cysteine residues on the second line, and two regions of homology with other protein Super Families (GFRD = Growth Factor Receptor Domain. EGF = EGF binding domain). The arrow indicates the exon containing the p.Cys75Ser mutation.