

G. S. Sellick · C. Longman · M. Brockington  
I. Mahjneh · L. Sagi · K. Bushby · H. Topaloğlu  
F. Muntoni · R. S. Houlston

## Localisation of merosin-positive congenital muscular dystrophy to chromosome 4p16.3

Received: 21 January 2005 / Accepted: 23 February 2005 / Published online: 11 May 2005  
© Springer-Verlag 2005

**Abstract** The congenital muscular dystrophies (CMD) are a heterogeneous group of autosomal recessive disorders, which present within the first 6 months of life with hypotonia, muscle weakness and contractures, associated with dystrophic changes on skeletal muscle biopsy. We have previously reported a large consanguineous family segregating merosin-positive congenital muscular dystrophy, in which involvement of known CMD loci was excluded. A genome-wide linkage search of the family conducted using microsatellite markers spaced at 10-Mb intervals failed to identify a disease locus. A second scan using a high-density SNP array, however, permitted a novel CMD locus on 4p16.3 to be identified (multipoint LOD score 3.4). Four additional consanguineous CMD families with a similar phenotype were evaluated for linkage to a 4.14-Mb interval on

4p16.3; however, none showed any evidence of linkage to the region. Our findings further illustrate the utility of highly informative SNP arrays compared with standard panels of microsatellite markers for the mapping of recessive disease loci.

### Introduction

The congenital muscular dystrophies (CMD) are a heterogeneous group of autosomal recessively inherited disorders that present at birth, or within the first 6 months of life, with muscle weakness, hypotonia and dystrophic changes on skeletal muscle biopsy (Dubowitz 1994). Many different forms of CMD are recognisable, based upon distinctive clinical, pathological and molecular abnormalities; however, patients remain who cannot be diagnosed as having one of the known CMD forms (Muntoni and Voit 2004).

We have studied one such family; a large, consanguineous Palestinian family with multiple members affected by CMD. Affected individuals had a neonatal presentation, poor muscle build, and normal or minimally elevated serum creatine kinase (CK) levels. They had a relatively static course and all patients who survived beyond early childhood acquired the ability to walk and all but one maintained this into adulthood. Linkage analysis has excluded involvement of mutations in the *LAMA2*, *ITGA7*, *FCMD* and *SEPNI* CMD genes (Mahjneh et al. 1999). The combination of clinical features in affected family members, and the exclusion of known CMD loci, allowed us to report this family previously as affected by a novel form of recessively inherited CMD (Mahjneh et al. 1999). Large families, such as the one we have identified provide a powerful means of identifying novel disease loci through autozygosity mapping.

Here we report the localisation of a novel locus for this form of merosin-positive CMD to the telomeric tip of chromosome 4p and compare the efficiency of per-

G. S. Sellick · R. S. Houlston  
Section of Cancer Genetics,  
Institute of Cancer Research,  
15 Cotswold Rd, Sutton,  
Surrey, UK

C. Longman · M. Brockington · L. Sagi · F. Muntoni (✉)  
Department of Paediatrics, Imperial College,  
Hammersmith Hospital,  
London, UK  
E-mail: f.muntoni@imperial.ac.uk  
Tel.: +44-20-8383-2480  
Fax: +44-20-8383-2473

C. Longman  
Department of Medical Genetics,  
Yorkhill Hospital, Glasgow, UK

I. Mahjneh  
University of Oulu and Pietarsaari Hospital,  
Pietarsaari, Finland

K. Bushby  
Institute of Human Genetics,  
University Newcastle upon Tyne,  
Newcastle, UK

H. Topaloğlu  
Department of Paediatric Neurology,  
Hacettepe Children's Hospital,  
Ankara, Turkey



**Table 1** Clinical manifestations of CMD in families 1–5: pattern of muscle involvement. *Mild* MRC grade 4, *moderate* MRC grade 3, *severe* MRC grade 2 or less. In family 1, the range of the observed features is indicated. Numbers of individuals denoted *in parentheses*

Family	Age at examination (years)	Facial weakness	Neck	Shoulder girdle	Forearm	Trunk	Pelvic girdle	Distal leg	Feet
1	23–38	Mild (6); Moderate (1)	Mild (3); Moderate (3); Severe (5)	Moderate (2); Severe (5)	Mild	Mild (2); severe (5)	Mild (2); Moderate (4); Severe (1)	Mild	Mild
2	12	Mild	Mild	Mild	Moderate	Moderate	Moderate	Moderate	Mild
3	12	Mild	Mild	Mild	Mild	Moderate	Moderate	Mild	Mild
4	14	None	Mild	Moderate	Mild	Mild	Moderate	Mild	Mild
5	10	Mild	Severe	Severe	Mild	Severe	Severe	Moderate	Mild

**Table 2** Clinical manifestations of CMD in families 1–5: signs and symptoms of patients. *Mild* MRC grade 4, *moderate* MRC grade 3, *severe* MRC grade 2 or less. In family 1, the range of the observed features is indicated. Numbers of individuals denoted *in parentheses*

Family	Sex of affected	Hypotonia at birth	Age at first walking (years)	Age at diagnosis (years)	Late contractures	Scoliosis	Creatine kinase level
1	2 females 5 males	Present (7/7)	2.5–5	2–15	Severe (1); Moderate (2); Mild (4)	Absent (2); Present (5)	Normal to fourfold increase
2	1 male	Present	Never	8	Severe	Absent	Threefold increase
3	1 female	Present	3	9	Absent	Present	Twofold increase
4	1 female	Present	3	11	Absent	Present	Threefold increase
5	1 female	Present	Never	4	Moderate	present	Twofold increase

age scan of family 1 was undertaken using the ABI Prism Linkage mapping sets 1 and 2 containing 400 markers. PCR reactions (5 µl total) contained 0.5 µl GeneAmp 10× PCR Gold buffer (Applied Biosystems, Warrington, UK), 2.5 mM MgCl<sub>2</sub> (final concentration) (Applied Biosystems), 0.125 U Amplitaq gold (Applied Biosystems), dNTP (1 mM each nucleotide, final concentration) (BD Biosystems), ABI Prism primer pair mix (1.25 µM each primer, final concentration), 2.5 µl DNA. PCR reactions were performed on a PE9600 Thermal-cycler (Applied Biosystems). PCR products were separated on a 5% denaturing gel (Amresco, Taipei, Taiwan) in an ABI 377A automated DNA sequencer (Applied Biosystems) and analysed using Genescan version 3.01 and Genotyper version 2.01 software (Applied Biosystems).

The second genome-wide linkage search of family 1 was conducted using the GeneChip Mapping 10k 131 *Xba*I Array containing 11,555 SNP markers. Genotypes were obtained by following the Affymetrix protocol (Affymetrix, Santa Clara, Calif., USA) (Matsuzaki et al. 2004). Briefly, 250 ng of genomic DNA was digested per sample with the restriction endonuclease *Xba*I for 2.5 h. Digested DNA was mixed with *Xba*I adapters and ligated using T4 DNA ligase for 2.5 h. Ligated DNA was added to four separate PCR reactions, cycled, pooled and purified to remove unincorporated ddNTPs. The purified PCR products were then fragmented and labelled with biotin-ddATP. Biotin labelled DNA fragments were hybridised to the arrays for 18 h in a standard Affymetrix 640 hybridisation oven. After hybridisation, arrays were washed, stained, and scanned

using an Affymetrix Fluidics Station F450 with images obtained by use of an Affymetrix GeneChip 3000 scanner. Affymetrix GCOS software (version 1.2) was used to obtain raw microarray feature intensities (RAS scores). RAS scores were processed using Affymetrix GDAS (version 3.0) software to derive SNP genotypes (Affymetrix).

Additional genotyping for chromosomes 13q (family 1) and 4p16.3 (families 1–5) was undertaken using fluorescence-labelled microsatellite markers (Invitrogen, UK) referenced to the UCSC Human Genome database (<http://genome.ucsc.edu>, July 2003 release). These analyses were performed on an ABI Prism 3100 genetic analyser with allele sizes determined by use of the ABI PRISM Genotyper software package (version 3.7) (Applied Biosystems).

#### Statistical analysis

A search for regions of microsatellite homozygosity was undertaken by inspection of genotypes in affected family members. For the SNP genome-wide scan a search for non-Mendelian inconsistencies and regions of haplotype sharing was initially undertaken using the statistical package STATA version 7.0 (Stata Corporation, Tex., USA).

Regions identified as being consistent with linkage were further analysed by multipoint linkage analyses using the Homoz program (Kruglyak et al. 1995). CMD was modelled as fully penetrant recessive disease with a population frequency of the disease allele set to 0.001.

As population frequency data for markers was unavailable for the specific ethnic group of the family, LOD scores were evaluated over a range of marker allele frequencies. The map order and distances between SNP and microsatellite markers was based on the UCSC Human Genome browser (<http://genome.ucsc.edu/>, July 2003 release).

### Sequencing

Mutational analysis of family 1 for spondin 2 (*SPON2*) and myosin regulatory light chain 5 (*MYL5*) variants was performed by direct sequencing. All exons and intron–exon boundaries were bi-directionally sequenced using ABI Big Dye Terminator chemistry (Applied Biosystems), according to manufacturer's instructions. Sequencing products were separated on a 5% denaturing gel (Amresco) on an ABI377A automated DNA sequencer (Applied Biosystems) and analysed using Sequence Analysis software version 2.01 (Applied Biosystems). PCR primers and conditions are available on request.

---

## Results

The initial genome-wide search of family 1 was conducted using the micro satellite markers within the ABI medium density linkage set. The linkage scan was based on genotypes obtained from typing the three affected individuals IV:1, IV:8 and IV:9 and the unaffected individuals III:1, III:2, III:3, III:4, III:5, III:8, IV:4, IV:5, IV:6, IV:7, IV:10 and IV:11. Of the 5,745 possible genotypes 95.6% of the markers were scoreable. Interrogation of marker genotypes did not permit a region of shared homozygosity amongst affected individuals to be unambiguously identified.

The second genome-wide linkage search was performed using the GeneChip Mapping 10k *Xba*I Array. Three of the affected family members, IV:1, IV:8 and IV:9 and four unaffected family members III:4, III:5, IV:4 and IV:7 were genotyped. The average genotype call rate obtained was 90.1% (range: 84.2–96.4%) providing data on  $10.4 \times 10^3$  genotypes per individual.

Two regions of homozygosity greater than 5 Mb were identified. The first mapped to a 6.3-Mb interval spanning the region 4p16.3–16.1 (defined by dbSNPs rs718429 to rs908015) and the second to a 6.2-Mb interval spanning 13q21.33–22.3 (defined by dbSNPs rs2147810 to rs1330907). Both regions of potential linkage were further evaluated using the microsatellite markers D4S2936, D4S3038, D4S1614, D4S3034, D4S412, D4S432, D4S2925 and D13S258, D13S800, D13S156, D13S792, D13S162, D13S782, D13S1306, respectively. In addition to genotyping IV:1, IV:8, IV:9, III:4, III:5, IV:4 and IV:7, individuals III:1, III:2, III:3, III:8, IV:5, IV:6, IV:10 and IV:11 were also genotyped for these markers (Fig. 1). Haplotype analysis confirmed

linkage of CMD to chromosome 4p in family 1 and excluded linkage to chromosome 13q. None of the unaffected family members were homozygous for the disease haplotype on chromosome 4p.

Combining data from both the SNP and microsatellite fine-mapping analyses permitted a 4.14-Mb interval of linkage to be delineated on chromosome 4p16.3, flanked centromerically by D4S432 (Fig. 1). The multipoint LOD score across the region of linkage was 3.4. Variation in marker allele frequencies did not significantly alter the lod score obtained.

Fifty-two RefSeq genes map to the region of linkage (UCSC Human Genome browser, <http://genome.ucsc.edu/>, release July 2003). Of these, 16 are predicted or hypothetical genes with little or no associated information regarding their biological function. At least two genes represent plausible candidates on the basis of their biology: *SPON2* and *MYL5*. We screened an affected family member for sequence changes in the coding regions and splice sites of both candidate genes. No obvious or potentially pathogenic changes were detected in either *SPON2* or *MYL5*.

To explore the possibility that the disease locus on chromosome 4p16.3 was causative of CMD in other consanguineous families accrued (families 2–5), we genotyped parents and all available family members using the microsatellite markers D4S2936, D4S3038, D4S1614, D4S3034, D4S412 and D4S2925. Haplotyping of marker genotypes provided no evidence of linkage of CMD to chromosome 4p16.3 in the four additional families.

---

## Discussion

It is becoming increasingly clear that identical CMD phenotypes can result from mutations in any one of several different genes. Currently, ten genes are known to cause CMD, and the resulting disorders can be broadly divided into three groups (Muntoni and Voit 2004). The largest group results from mutations in genes encoding structural proteins. The commonest form, merosin-deficient CMD (MDC1A), results from mutations in *LAMA2*, encoding the laminin- $\alpha$ 2 chain of laminin-2 (merosin), a key component of the muscle cell basement membrane (Helbling-Leclerc et al. 1995). Another common form of CMD, Ullrich syndrome, results from mutations in the collagen VI chain genes, *COL6A1* (Lampe et al. 2005), *COL6A2* (Camacho et al. 2001; Higuchi et al. 2001) or *COL6A3* (Demir et al. 2002), which comprise the extra cellular matrix protein, collagen VI. Involvement of each of these genes was excluded by linkage analysis in our family. The second group, collectively termed the dystroglycanopathies, result from mutations in genes encoding known or putative glycosyltransferase enzymes involved in glycosylation of the sarcolemmal protein  $\alpha$ -dystroglycan. These include Fukuyama CMD, Walker Warburg syndrome, muscle-eye-brain disease, MDC1C and

MDC1D, due to mutations in *FCMD* (Kobayashi et al. 1998), *POMT1* (Beltran-Valero de Bernabe et al. 2002), *POMGnT1* (Yoshida et al. 2001), *FKRP* (Brockington et al. 2001) and *LARGE* (Longman et al. 2003), respectively. Another dystroglycanopathy locus (*MDC1B*) maps to chromosome 1q42, but the gene responsible has not been identified (Brockington et al. 2000). These patients rarely achieve independent ambulation and, with the exception of MDC1B and MDC1C cases, have mental retardation and a neuronal migration defect on brain magnetic resonance imaging, thus are clinically unlike our cases. These dystroglycanopathy patients also show abnormal labelling of  $\alpha$ -dystroglycan in muscle, but this was not examined in our cases. However, a secondary deficiency of laminin  $\alpha 2$  is usually also observed, whereas laminin  $\alpha 2$  was normal on both immunolabelling and immunoblotting in the family we describe. Furthermore, the dystroglycanopathies are invariably characterised by severe and progressive muscle damage, with markedly elevated serum CK, features not present in the family we have studied. Finally, mutations in *SEPN1*, encoding an endoplasmic reticulum resident selenoprotein of unknown function, result in CMD with spinal rigidity (RSMD1) (Moghadaszadeh et al. 2001); however, involvement of *SEPN1* was excluded by linkage analysis in our family. Although there are several other clinically recognisable CMD phenotypes in which the molecular defect is unknown (Muntoni and Voit 2004), our patients do not resemble any of these forms. We therefore conclude that they are affected by a novel form of CMD.

The underlying gene defect in family 1 remains unknown. Although we cannot exclude a contiguous-gene syndrome, the monosystemic nature of the condition argues against this possibility. In addition, any such deletion would have to be < 400 kb in size, as genotype signatures were obtained from all SNP and microsatellite markers mapping to the region of linkage. A number of genes mapping to chromosome 4p16.3 represent plausible candidates on the basis of either muscle expression and/or implied biological function of the expressed protein. Of these, two were screened for mutations as they were considered attractive candidates. *SPON2* is expressed in skeletal muscle and predicted to encode an extra cellular matrix protein (Manda et al. 1999). *MYL5* codes for a regulatory light chain of myosin and is expressed in skeletal muscle (Collins et al. 1992). Mutations in the related *MYL2*, expressed in heart, result in familial hypertrophic cardiomyopathy and skeletal myopathy (Poetter et al. 1996). However, no pathogenic sequences changes were detected in the coding region of these genes, making it unlikely that defects in either *SPON2* or *MYL5* underlie CMD in our family. Joint laxity, a characteristic feature of Ullrich syndrome, was observed in our family. Genes encoding collagens, other than type 6A, would therefore be good functional candidates; however, no such genes localise to chromosome 4p16.3. The affected individuals in the other families' studied had a clinical phenotype similar

to that observed in affected individuals from family 1, inviting speculation that the same disease locus might be involved. Linkage data, however, excluded involvement of the disease locus on chromosome 4p16.3 in families 2–5, suggesting further genetic heterogeneity in CMD within this phenotype.

As shown by the analysis of family 1, autozygosity mapping provides a powerful means of identifying recessive disease loci. Conventionally, genome-wide linkage searches have been conducted using microsatellite markers. Using microsatellite markers spaced at 10-Mb intervals for an initial scan has proven to be effective in identifying a region of homozygosity amongst affecteds; however, it is often necessary to undertake additional genotyping. This can be highly time consuming and labour intensive. Our initial genome-wide scan using microsatellites proved ineffective in identifying a region of potential linkage due to a lack of informative markers in the linked region. Rather than attempt to interrogate numerous genomic regions of low information content with additional microsatellite markers our recent experience suggested that using a high-density SNP array might prove to be more productive (Sellick et al. 2004). This subsequent analysis led quickly to the identification of a novel CMD locus. Therefore, we conclude that previous genome-wide linkage searches of consanguineous families based on microsatellite markers, which failed to identify a disease locus merit repeating using high-density SNP-based arrays where panel information content is substantially increased.

**Acknowledgments** The authors wish to thank the families and their physicians for their co-operation. Gabrielle Sellick was in receipt of a Postdoctoral Research Fellowship from Leukaemia Research. A grant from the Muscular Dystrophy Campaign to F.M. is gratefully acknowledged. Conflict of interest: none declared.

## References

- Beltran-Valero de Bernabe D, Currier S, Steinbrecher A, Celli J et al (2002) Mutations in the O-mannosyltransferase gene *POMT1* give rise to the severe neuronal migration disorder Walker–Warburg syndrome. *Am J Hum Genet* 71:1033–1043
- Brockington M, Sewry CA, Herrmann R, Naom I et al (2000) Assignment of a form of congenital muscular dystrophy with secondary merosin deficiency to chromosome 1q42. *Am J Hum Genet* 66:428–435
- Brockington M, Blake DJ, Prandini P, Brown SC et al (2001) Mutations in the fukutin-related protein gene (*FKRP*) cause a form of congenital muscular dystrophy with secondary laminin alpha2 deficiency and abnormal glycosylation of alpha-dystroglycan. *Am J Hum Genet* 69:1198–1209
- Camacho VO, Bertini E, Zhang RZ, Petrini S et al (2001) Ullrich scleroatonic muscular dystrophy is caused by recessive mutations in collagen type VI. *Proc Natl Acad Sci USA* 98:7516–7521
- Collins C, Schappert K, Hayden MR (1992) The genomic organisation of a novel regulatory myosin light chain gene (*MYL5*) that maps to chromosome 4p16.3 and shows different patterns of expression between primates. *Hum Mol Gen* 1:727–733
- Demir E, Sabatelli P, Allamand V, Ferreiro A et al (2002) Mutations in *COL6A3* cause severe and mild phenotypes of Ullrich congenital muscular dystrophy. *Am J Hum Genet* 70:1446–1458

- Dubowitz V (1994) 22nd ENMC sponsored workshop on congenital muscular dystrophy held in Baarn, The Netherlands, 14–16 May 1993. *Neuromuscul Disord* 4:75–81
- Helbling-Leclerc A, Zhang X, Topaloglu H, Cruaud C et al (1995) Mutations in the laminin alpha 2-chain gene (*LAMA2*) cause merosin-deficient congenital muscular dystrophy. *Nat Genet* 11:216–218
- Higuchi I, Shiraishi T, Hashiguchi T, Suehara M et al (2001) Frameshift mutation in the collagen VI gene causes Ullrich's disease. *Ann Neurol* 50:261–265
- Kobayashi K, Nakahori Y, Miyake M, Matsumura K et al (1998) An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. *Nature* 394:388–392
- Kruglyak L, Daly MJ, Lander ES (1995) Rapid multipoint linkage analysis of recessive traits in nuclear families, including homozygosity mapping. *Am J Hum Genet* 56:519–527
- Lampe A, Dunn DM, von Niederhauserm AC, Hamil C et al (2005) Automated genomic sequence analysis of the three collagen VI genes: applications to Ullrich congenital muscular dystrophy and Bethlem myopathy. *J Med Gen* 42:108–120
- Longman C, Brockington M, Torelli S, Jimenez-Mallebrera C et al (2003) Mutations in the human *LARGE* gene cause MDC1D, a novel form of congenital muscular dystrophy with severe mental retardation and abnormal glycosylation of alpha-dystroglycan. *Hum Mol Genet* 12:2853–2861
- Mahjneh I, Vannelli G, Bushby K, Marconi GP (1992) A large inbred Palestinian family with two forms of muscular dystrophy. *Neuromuscul Disord* 2:277–283
- Mahjneh I, Bushby K, Anderson L, Muntoni F et al (1999) Merosin-positive congenital muscular dystrophy: a large inbred family. *Neuropediatrics* 30:22–28
- Manda R, Kohno T, Matsuno Y, Takenoshita S et al (1999) Identification of genes (*SPON2* and *C20orf2*) differentially expressed between cancerous and noncancerous lung cells by mRNA differential display. *Genomics* 61:5–14
- Matsuzaki H, Loi H, Dong S, Tsai YY et al (2004) Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. *Genome Res* 14:414–425
- Moghadaszadeh B, Petit N, Jaillard C, Brockington M et al (2001) Mutations in *SEPN1* cause congenital muscular dystrophy with spinal rigidity and restrictive respiratory syndrome. *Nat Genet* 29:17–18
- Muntoni F, Voit T (2004) The congenital muscular dystrophies in 2004: a century of exciting progress. *Neuromuscul Disord* 14:635–649
- Poetter K, Jiang H, Hassanzadeh S, Master et al (1996) Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. *Nat Genet* 13:63–69
- Sellick GS, Longman C, Tolmie J, Newbury-Ecob R et al (2004) Genome-wide linkage searches for Mendelian disease loci can be efficiently conducted using high-density SNP genotyping arrays. *Nucleic Acids Res* 32:e164
- Yoshida A, Kobayashi K, Manya H, Taniguchi K et al (2001) Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Dev Cell* 1:717–724