## ARTICLE

# Congenital cataracts facial dysmorphism neuropathy (CCFDN) syndrome: a novel developmental disorder in Gypsies maps to 18qter

Dora Angelicheva<sup>1</sup>, Ivailo Turnev<sup>2</sup>, Danielle Dye<sup>1</sup>, David Chandler<sup>1</sup>, PK Thomas<sup>3</sup> and Luba Kalaydjieva<sup>1</sup>

<sup>1</sup>Centre for Human Genetics, Edith Cowan University, Perth, Australia <sup>2</sup>Department of Neurology, Medical University, Sofia, Bulgaria <sup>3</sup>Department of Clinical Neurosciences, Royal Free and University College School of Medicine, London, UK

We have identified a novel developmental disorder with complex phenotypic characteristics involving primarily the nervous system, which appears to be common in a specific Gypsy group in Bulgaria. We propose to refer to the syndrome as congenital cataracts facial dysmorphism neuropathy (CCFDN). We have assigned the disease locus to the telomeric region of chromosome 18q. Linkage disequilibrium and highly conserved haplotypes suggest genetic homogeneity and founder effect. CCFDN co-localises with an EST which shows high homology to a conserved *Drosophila* gene involved in the regulation of nervous system development in vertebrates.

Keywords: developmental disorder; hypomyelinating neuropathy; gene mapping; genetic isolate; Gypsies

## Introduction

Our previous study of hereditary motor and sensory neuropathy, Lom (HMSNL), a novel autosomal recessive demyelinating neuropathy associated with deafness, was the first to indicate that, despite their geographic dispersal and social and linguistic adaptability, some Gypsy groups have remained genetically restricted and display the characteristics of isolated founder populations.<sup>1,2</sup> During follow-up studies aimed at the identification of additional HMSNL patients in Bulgaria, we have come across another novel disorder which presents as a complex developmental abnormality involving primarily the nervous system. We have named this disorder the Congenital Cataracts Facial Dysmorphism Neuropathy (CCFDN) syndrome, to emphasise some of its salient phenotypic characteristics. CCFDN was initially identified in a large Gypsy kindred where five patients from two different branches of the pedigree displayed all the clinical features of Hereditary Motor and Sensory Neuropathy-Lom, whereas two female siblings from a third branch presented with a novel complex phenotype, also including a peripheral neuropathy. A subsequent search for patients with the same phenotype has resulted, within less than one year, in the identification of 50 affected

Correspondence: Dr Luba Kalaydjieva, Centre for Human Genetics, Edith Cowan University Joondalup Campus, Perth 6027 WA, Australia. Tel: +61894005808; Fax: +61894005851; E-mail: L.Kalaydjieva@cowan.edu.au

Received 18 September 1998; revised 23 November 1998; accepted 28 December 1998

individuals, suggesting that the disorder may be common among specific Gypsy groups. Here we report on the localisation of the *CCFDN* gene.

## **Subjects and Methods**

## Case Ascertainment, Pedigree Information and Gypsy Group Identity

Index patients were ascertained by searching hospital records, particularly of paediatric neurology and electrophysiology units, for a combination of peripheral neuropathy and congenital cataracts. Further affected relatives were traced through extensive interviews with key informants in each kindred. In addition, new cases were identified via extensive field work in Gypsy communities known to belong to the high-risk Gypsy group. To date, 19 affected families with a total of 50 CCFDN patients have been diagnosed.

Pedigree information was collected during interviews with informants in each family, mainly from the older generations, and verified by cross-checking the data provided by members of different branches of the family.

The aims of the study were explained in detail and discussed with each family and informed consent for participation was obtained.

#### Phenotype Assessment

All 50 affected individuals underwent physical and neurological examinations in their homes and at least one patient from each family was hospitalised for further investigations. Nerve conduction was examined in the upper and lower limbs in 23 patients using standard procedures. Standard electroencephalographic records were performed in 12 individuals. Magnetic resonance imaging of the brain and spinal cord was undertaken in 16 patients. Neuropathological studies were conducted on four sural nerve biopsy samples. Ophthalmological examinations were performed in 37 cases. Cognitive function was formally assessed in 20 patients. Endocrine function was evaluated in 15 patients.

#### Gene Mapping Strategy

The gene mapping study included 13 CCFDN families with a total of 86 members (35 affected and 51 unaffected). The initial stage of the study included a genome scan on 16 affected individuals from seven families, aiming to exclude a large proportion of the genome and identify chromosomal regions where excess sharing IBS occurred in affected relatives. Regions of interest were studied further by genotyping all members of these seven families, including parents and unaffected sibs. In the final stage of the study, all affected families were analysed for the candidate region.

The initial screen included the four affected individuals in family D who are closely related to each other in different ways (Figure 1). Previous experience with genetic studies in Gypsies<sup>1</sup> indicates that a high frequency and an old age of the mutation may result in diverse disease haplotypes segregating within a single kindred. Therefore we did not postulate that a single polymorphic allele should be shared between all affected individuals in family D but rather scored each locus for allele sharing in pairs D5/D2; D2/D8 and D5/D9. Since consanguinity in family D as well as the possible high level of background inbreeding could be expected to result in false positive findings, stage 1 also included the analysis of six affected sib pairs scored as sharing zero, one or two alleles IBS. The scores for family D and for the sib pairs were plotted for each chromosome and the genotyping data were also used in multipoint Lod score calculations.

The ABI PRISM<sup>™</sup> Linkage Mapping Set panels (Versions 1 and 2) (Perkin-Elmer, Foster City, USA) was used for the genome scan. Additional markers on 18qter were selected from the Généthon map.<sup>3</sup> Two-point lod score analysis was conducted using FASTLINK,<sup>4</sup> under an autosomal recessive model with 100% penetrance and three different gene frequencies (0.01, 0.05 and 0.10). Multipoint linkage analysis was performed using GENEHUNTER,<sup>5</sup> after breaking the loops in the consanguinous families. The marker order and map distances were as provided in the ABI map. Haplotypes were constructed manually and analysed for conserved regions and evidence of historical recombinations. Linkage disequilibrium was calculated using the transmission disequilibrium test.<sup>6</sup>

#### Genotyping of Microsatellite Markers

DNA was extracted using the salting-out method<sup>7</sup> from 5-10 ml of venous blood in 59 cases. Samples from 26 individuals were obtained as dried blood spots taken on FTA-Gene Guard System (Life Technologies, Rockville, USA) filter paper cards and DNA was isolated following the manufacturer's protocols.

ABI panel markers were amplified in a 5µl final volume containing 5–10 ng genomic DNA,  $0.2 \mu$ M of each primer, 250µM of each dNTP,  $1.5 \,$ mM MgCl<sub>2</sub> 1 × PCR buffer II (Perkin Elmer N808-0010) and 0.2 units of Amplitaq<sup>TM</sup> DNA polymerase (Perkin Elmer N808-0101). The PCR products were electrophoresed on an ABI 373 XL DNA Sequencer in a 24 cm well to read, 6%, 19:1 bis/polyacrylamide gel at 1400V for 7h. Allele analysis was conducted using ABI GENO-TYPER<sup>TM</sup> (Version 2) software.

Additional markers on 18q were analysed using radioactive labelling, by reducing dNTP concentration to 25  $\mu$ M each and adding 0.2  $\mu$ Ci of 300 Ci/mmol  $\alpha$ -<sup>32</sup>P dCTP (Bresatec, Adelaide, Australia) to each PCR reaction. The PCR products were electrophoresed on 6%, 19:1 bis/polyacrylamide gel at 1400V for 2–3 h. Autoradiography was for 4–48 h. Alleles were scored manually.

## Results

#### The CCFDN Phenotype

The 50 affected individuals ranged in age between 8 months and 40 years. The disorder was first recognised in infancy by the presence of congenital cataracts and microcorneae. Motor and intellectual development were delayed. Facial dysmorphism became evident during childhood, characterised by a prominent midface, thickening of the perioral tissues, forwardly directed anterior dentition and hypognathism.

A progressive distally accentuated, predominantly motor peripheral neuropathy affecting first the lower and then the upper limbs developed during childhood and adolescence and was associated with skeletal



**Figure 1** Family D was included, together with six sib pairs, in the initial genome scan which searched for sharing of marker alleles between individuals D2/D5, D2/D8 and D5/D9. The figure presents the genotypes of the entire pedigree for the markers in the 18q23–qter region

deformities. Motor and sensory conduction studies showed slowing into the demyelinating range. Nerve biopsy examination indicated generalised hypomyelination superimposed upon which were demyelination and axonal degeneration in older subjects. Central nervous system involvement was evidenced by a mild non-progressive cognitive deficit, accompanied in some patients by extensor plantar responses, mild chorea, upper limb postural tremor and mild ataxia. Magnetic resonance imaging demonstrated cerebral and spinal cord atrophy and occasional focal lesions.

Associated non-neural features included short stature and hypogonadism with secondary amenorrhea in female subjects. Hormonal investigation showed low normal growth hormone values and reduced testosterone and estradiol levels. A detailed description of the clinical and neuropathological findings will be presented elsewhere, Turnev I *et al* (personal communications).<sup>8</sup>

#### **CCFDN** Families

Thus far, the search for CCFDN families has covered primarily the northern part of Bulgaria, between the Danube river and the Balkan mountain range. Nineteen extended families with the disorder have been identified, mostly originating from an endogamous Gypsy group known as the Kopanari (an ethnonym derived from the traditional trade of the group where *kopana* is an old Bulgarian word for a wooden bowl).<sup>9</sup> The Kopanari migrated to Bulgaria from the Wallachian Kingdom, which was part of present-day Rumania, and from Moldova. The migrations are believed to have

562

groups. Of the 19 CCFDN families identified to date, 13 were included in the gene mapping study. Consanguinity was reported in five families, with close consanguinity in three of these. Information on distant relationships within the large kindreds was readily provided in great detail despite the fact that the different branches of the large kindreds were scattered over several hundred kilometers. However, none of the CCFDN families was aware of the existence of the other affected kindreds and no common ancestor could be traced in the last 6–8 generations.

Autosomal recessive inheritance was suggested by the pedigree structure. All parents were clinically unaffected. Electrophysiological studies demonstrated normal nerve conduction velocity.

## **Gene Mapping**

#### Stage 1 Genome Scan

Genome screening showed sharing between the affected relatives in family D for 40% of the markers. In the different sib pairs, IBS sharing of both polymorphic alleles ranged between 34% and 44% of loci. This stage identified 17 genomic regions with evidence of two or fewer recombinations in the overall group. Only one region, namely 18q23–qter, gave consistently positive results for three adjacent markers spanning a distance of about 11 cM on the ABI map. The affected individuals in family D shared alleles for D18S469, D18S462 and D18S70 (Figure 1). Sharing of both alleles was found in 5 of 6 affected sib pairs for markers D18S469 and D18S462 and in 6 of 6 sib pairs for marker D18S70. All 16 affected individuals were homozygous

for the same allele of marker D18S70. These findings pointed to the telomeric part of 18q as the most likely location of the *CCFDN* gene.

#### The Localisation of the CCFDN Gene

Genotyping the complete set of 13 CCFDN families with a total of 86 members for markers D18S469, D18S462 and D18S70 produced a maximum 2-point lod score of 7.64 at D18S70, recombination distance 0.00 (Table 1).

A denser map of the 18q region was constructed by adding microsatellites D18S1115, D18S461, D18S1122 and D18S1141, placed on the Généthon map<sup>3</sup> between D18S462 and D18S70, and D18S1268 which, according to the HCLC map, is located between D18S70 and the telomere. The entire set of CCFDN families was typed for the five newly added markers. The two point lod scores are presented in Table 1. The highest value of 9.02 was at D18S1141, recombination fraction 0.00. Changing the gene frequency did not substantially affect the results. The multipoint lod scores peaked at D18S70, with a value of 12.

Haplotypes were manually constructed for the eight markers in the 18q23-qter region in the order D18S462-D18S1115-D18S461-D18S1122-

D1141–D18S70–D18S1268. A recombination in family WA placed the gene telomeric to D18S461 and another, in an unaffected sibling in family WS, helped to narrow down the region further and placed the gene telomeric to D18S1141 (not shown).

CCFDN alleles presented a highly conserved haplotype in the region D18S1141–D18S70–D18S1268 (Figure 2). With the exception of a single chromosome, all patients were homozygous for the same allele of

**Table 1** Two-point Lod scores for markers in the 18q23-qter region obtained under an autosomal recessive model with 100%penetrance and gene frequency 0.05

Marker	Recombination fraction								
	0.0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4
D18S469	_	2.14	2.18	1.91	1.54	1.14	0.77	0.45	0.21
D18S462	_	3.91	3.51	2.88	2.20	1.56	1.01	0.57	0.25
D18S1115	2.24	1.89	1.53	1.19	0.89	0.62	0.40	0.22	0.10
D18S461	-	5.83	5.04	4.08	3.12	2.23	1.45	0.82	0.37
D18S1122	6.00	5.31	4.43	3.52	2.66	1.88	1.22	0.70	0.32
D18S1141	9.02	7.60	6.19	4.85	3.62	2.54	1.64	0.93	0.43
D18S70	7.64	6.42	5.24	4.12	3.10	2.20	1.44	0.83	0.39
D18S1268	5.47	4.57	3.68	2.84	2.09	1.44	0.91	0.51	0.23

CCFDN mapping in Gypsies D Angelicheva *et al* 





**Figure 2** Marker haplotypes of 70 CCFDN chromosomes in the 18qter region. Marker order and distances are as shown in the Généthon map. Asterisks indicate possible microsatellite mutations

marker D18S70 which displayed strong linkage disequilibrium with the disease (see (Figure 2) for *P*-values obtained in the TD test). The CCFDN chromosome which carried a different D18S70 allele had the conserved haplotype for the remaining markers (bottom line in Figure 2), indicating a possible microsatellite mutation in D18S70. These findings point to genetic homogeneity and a single founder mutation carried by all CCFDN alleles. Several historical recombination events involving either D18S1141 and markers centromeric to it or D18S1268, place the gene between these two loci, in the interval around D18S70.

## Discussion

#### The CCFDN Phenotype

No previous description of a disorder fitting the clinical manifestations of CCFDN could be identified in the literature or in the London Dysmorphology Database.

**5**64

Several conditions of autosomal recessive inheritance partly overlap the CCFDN phenotype, eg the Marinesco–Sjögren syndrome;<sup>10</sup> mental retardation, congenital cataracts, sensorineural deafness, spinocerebellar ataxia and reduced MNCV;<sup>6</sup> and an Indian family with hypertrophic neuropathy and cataracts.<sup>12</sup> None display the whole spectrum of symptoms which characterise CCFDN. We can therefore conclude that the Congenital Cataracts Facial Dysmorphism Neuropathy syndrome is a single gene disorder of autosomal recessive inheritance which has not been described previously.

#### Gene Mapping in Gypsies

CCFDN is the second novel nosological entity identified since the beginning of our studies of inherited diseases in Balkan Gypsies in 1993. In both disorders, the genotyping results strongly suggest genetic homogeneity and recent founder effect. The emerging pattern resembles that of the 'Finnish disease heritage<sup>113</sup> and places the Gypsies among the most prominent genetic isolates of Europe where novel disorders can be found and founder effect and linkage disequilibrium can be exploited in genetic research. The Gypsies form a complex population with marked internal stratification,<sup>1,5,14</sup>, and awareness of its existence is important for the identification of patients with specific disorders and for defining high-risk groups where targeted prevention can be of immediate benefit.

Given the current resolution of the standard genetic maps, homozygosity mapping does not seem to be appropriate, since an old mutation and a complex social and demographic history of the affected Gypsy group could produce misleading results. However, with the advent of new technologies and higher resolution maps,<sup>15</sup> it will become a powerful strategy for localising disease genes in isolated Gypsy populations.

#### Candidate Genes

The multiple features of CCFDN suggest a disorder in early development. In the peripheral nerve, our studies have shown that the Schwann cell fails to produce myelin of adequate thickness. It is impossible at this stage to predict whether hypomyelination in CCFDN results from impaired axonal signalling or from a defect in the Schwann cell response. The nature and sequence of abnormalities and the causes of variation in the clinical manifestations of central nervous system involvement also remain unclear. Endocrine dysfunction may be secondary to hypothalamic involvement or may be primary, if the gene is normally expressed in the gonads. The early manifestation of cataracts and microcornea indicate a primary defect in the development of the eye. Therefore the *CCFDN* gene can be predicted to be involved in early development and/or differentiation of a number of cell types. It could thus be a member of growth factor, growth factor receptor or transcription regulator gene families. Several genes have been mapped close to the CCFDN region. Myelin basic protein (MBP) is a structural myelin component and forms part of a complex transcriptional unit.<sup>16</sup> Closely linked to MBP is the galanin receptor gene. Galanin is a neuromodulator present in the brain and the hypothalamopituitary axis and is a potent stimulator of growth hormone secretion. In view of the low stature observed in CCFDN patients, the recent finding of a high frequency of deletions involving the galanin receptor gene in patients with growth retardation may be of interest.<sup>17</sup> One could speculate on the possible involvement of these genes in CCFDN; however, none has an attributed role in developmental regulation that could explain the pleomorphic CCFDN phenotype. A search of expressed sequence tags (ESTs) in the public databases has identified an EST placed in the 18qter region, with a high degree of homology to msal, the mouse homolog of the Drosophila spalt gene. Spalt is a member of a distinct class of zinc-finger proteins.<sup>18</sup> Its homologs have been shown to be essential for neuronal differentiation in the mouse, Xenopus and the Medaka fish.<sup>20,21</sup> In some vertebrate species, the gene is also expressed in other tissues such as the gonads and the developing eye. In humans, two spalt homologs, designated Hsal1 and Hsal2, have been mapped to 16q12.1 and 14q11.1–12.1.<sup>22</sup> Hsal1 has been implicated in a rare autosomal dominant malformation syndrome presenting with anal, renal, limb and ear abnormalities.<sup>23</sup> Interestingly, the EST mapped to 18q shows higher homology to the mouse, Xenopus and Medaka fish spalt genes than to the human Hsal1 and Hsal2 genes. It is therefore probable that the human sal gene family includes additional members and that the gene on chromosome 18q is the closest structural and functional homolog of the spalt genes in other vertebrates. A characterisation of this novel gene and a study of its involvement in CCFDN are currently in progress.

## Acknowledgements

This study was funded by Edith Cowan University and by the Association Française contre les Myopathies. The authors wish to thank Professor R Winter, Institute of Child Health, London, UK, and Professor N Risch, Stanford University School of Medicine, USA, for helpful discussions, the clinical team at the Medical University in Sofia and Drs RHM King and B Youl for the investigation of the patients and A Savov and O Kamenov for technical help. PKT wishes to acknowledge support from the Wellcome Trust, and IT from the Open Society Foundation.

## References

- 1 Kalaydjieva L, Hallmayer J, Chandler D *et al*: Gene mapping in Gypsies identifies a novel demyelinating neuropathy on chromosome 8q24. *Nat Genet* 1996; **14**: 214–217.
- 2 Kalaydjieva L, Nikolova A, Turnev I *et al*: Hereditary motor and sensory neuropathy Lom: a novel demyelinating neuropathy associated with deafness in Gypsies: clinical, electrophysiological and nerve biopsy findings. *Brain* 1998; **121**: 399–408.
- 3 Dib C, Fauré S, Fizames C *et al*: A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 1996; **380**: 152–154.
- 4 Lathrop GM, Lalouel J-M: Easy calculations of lod scores and genetic risks on small computers. *Am J Hum Genet* 1984; **36**: 460–465.
- 5 Kruglyak L, Daly MJ, Lander ES: Rapid multipoint linkage analysis of recessive traits in nuclear families, including homozygosity mapping. *Am J Hum Genet* 1995; 56: 519–527.
- 6 Sham PC, Curtis D: An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Ann Hum Genet* 1995; **59**: 323–336.
- 7 Miller SA: A simple salting out procedure for extracting DNA from human nucleated cells. *NAR* 1988; **16**: 1215.
- 8 Turnev I, Kalaydjieva L, Youl B *et al*: The congenital cataract facial dysmorphism neuropathy (CCFDN) syndrome, a novel complex disease in Balkan gypsies: clinical and electrophysiological observations. *Annals of Neurology* 1999; in press.
- 9 Marushiakova E, Popov V: Gypsies (Roma) in Bulgaria. Studien zur Tsiganologie und Folkloristik. Peter Lang: Frankfurt am Main, 1997, vol. 18.
- 10 Sjögren T: Hereditary congenital spinocerebellar ataxia accompanied by congenital cataract and oligophrenia. *Confin Neurol* 1950; **10**: 293–308.

- 11 Beeger JH, Scholte FA, Van Essen AJ: Two sisters with mental retardation, cataract, ataxia, progressive hearing loss and polyneuropathy: *J Med Genet* 1991; **28**: 284–285.
- 12 Gold GN, Hogenhuis LA: Hypertrophic interstitial neuropathy and cataracts. *Neurology* 1968; **18**: 526–533.
- 13 de la Chapelle A: Disease gene mapping in isolated human populations: the example of Finland. *J Med Genet* 1993; **30**: 857–865.
- 14 Liègeois J-P: *Roma, Gypsies, Travellers.* Council of Europe Press: Strasbourg, 1994.
- 15 Kruglyak L: The use of a genetic map of biallelic markers in linkage studies. *Nat Genet* 1997; **17**: 21–24.
- 16 Pribyl TM, Campagnoni CW, Kampf K: The human myelin basic protein is included within a 179-kilobase transcription unit: expression in the immune and central nervous system. *PNAS* 1993; **90**: 10695–10699.
- 17 Cody JD, Hale DE, Brkanac Z, Kaye C, Leach RJ: Growth hormone insufficiency associated with haploinsufficiency at 18q23. Am J Med Genet 1997; 71:420–423.
- 18 Kuhnlein RP, Frommer G, Friedreich M, Gonzalez-Gaitan M et al: Spalt encodes an evolutionary conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the Drosophila embryo. EMBO J 1994; 13: 168–179.
- 19 Ott T, Kaestner KH, Monaghan AP, Schutz G: The mouse homolog of the region specific homeotic gene *spalt* of *Drosophila* is expressed in the developing nervous system and in mesoderm-derived structures. *Mech Dev* 1996; 56: 117–128.
- 20 Hollemann T, Schuh R, Pieler T, Stick R: Xenopus Xsal1, a vertebrate homolog of the region specific homeotic gene *spalt* of *Drosophila*. *Mech Dev* 1996; **55**: 19–32.
- 21 Koster R, Stick R, Loosli F, Wittbrodt J: Medaka *spalt* acts as a target gene of hedgehog signalling. *Development* 1997; **55**: 19–32.
- 22 Kohlhase J, Schuh R, Dowe G *et al*: Isolation, characterization and organ-specific expression of two novel human zinc finger genes related to the Drosophila gene *spalt*. *Genomics* 1996; **38**: 291–298.
- 23 Kohlhase J, Wischermann A, Reichenbach H, Froster U, Engel W: Mutations in the SAL1 putative transcription factor gene cause Townes–Brocks syndrome. Nat Genet 1998; 18: 81–83.

566