University of Siena
Department of Biotechnology

Ph.D in Medical Genetics

Advances in Cohen syndrome diagnosis by MLPA and NGS

Veronica Parri

Tutor: Prof. Francesca Ariani

Supervisor: Prof. Alessandra Renieri

Doctoral school in Oncology and Genetics

Academic year 2010-2011

Cicle XXIII

Abstract

Cohen syndrome is a rare autosomal recessive disorder characterized by intellectual disability, typical facial features, hypotonia, pigmentary retinopathy, myopia and intermittent neutropenia. In 2003, mutations in *COH1* gene (8q22) were identified as causative of Cohen syndrome. *COH1* consists of 62 exons and encodes for a peripheral Golgi protein probably involved in intracellular vesicle mediated sorting and transport of proteins. In our laboratory, traditional methods of point mutation analysis consists of denaturing high performance liquid chromatography (DHPLC) followed by automatic sequencing. The screening of a group of 96 patients with a hypothetical diagnosis of Cohen syndrome revealed 21 different point mutations, including frameshift, splicing and missense mutations. In some patients, only one or no pathogenic variants were found by traditional techniques. Since deletions have been reported as a cause of Cohen syndrome, it was possible that large rearrangements could account for missed *COH1* mutations. Then, Multiple Ligation-dependent Probe Amplification (MLPA) was used to screen for *COH1* large rearrangements in a group of 14 patients with a phenotype strongly suggestive of Cohen syndrome. Using two kits containing probes for 60/62 exons of *COH1*, 11 deletions and 4 duplications were disclosed. Three patients shared the same deletion spanning exons 6-16 that was also reported in a large Greek consanguineous family. Haplotype analysis suggested that the recurrent deletion is due to a founder effect in the Mediterranean area. Since duplications has never been reported before in Cohen syndrome patients, Long Template PCR and automatic sequencing were used to characterize one of them, spanning exons 57-60. Determination of the initial breakpoint suggested that the duplication leads to a frameshift and a premature truncation of the protein. Finally, since the traditional

mutational analysis of the gene is expensive and time-consuming, we set up the conditions for the application of Next Generation Sequencing (NGS) to Cohen syndrome diagnosis and we simultaneously screened all the *COH1* exons in the first patient. In conclusion, incorporation of MLPA and NGS techniques demonstrated to increase the percentage of mutated *COH1* alleles and to save time and costs respect to the other methods presently reported.

INDEX

1. Introduction

1. INTRODUCTION

1.1 Cohen syndrome

Cohen syndrome (MIM 216550) is a rare autosomal recessive disorder first described in 1973 by Michael Cohen and collaborators. They observed two sibs and an unrelated patient with a previously unrecognised pattern of anomalies, including intellectual disability, facial dysmorphisms, hypotonia, obesity and ocular anomalies[1]. These patients were diagnosed with a newly recognized syndrome that Carey and Hall established as a clinical entity some years later [2].

In the Finnish population Cohen syndrome phenotype has been initially described as homogeneous [3], consisting of nonprogressive mild to severe psychomotor retardation, motor clumsiness, microcephaly, characteristic facial features, childhood hypotonia and joint laxity, progressive retinochoroidal dystrophy, myopia, intermittent isolated neutropenia, and a cheerful disposition. Characteristic facial features include high-arched or wave-shaped eyelids, short philtrum, thick hair, and low hairline [1-4] (Fig. 1A). In non-Finnish patients, the clinical phenotype is instead highly variable, lacking some of the traditional characteristic findings or showing additional features [5-7].

In 2003, the identification of *COH1* as the gene responsible of Cohen syndrome revealed that the highly homogeneous phenotype in Finnish patients was due to a strong founder effect. In fact, molecular analysis of *COH1* discovered an ancestral mutation responsible of the majority of cases [8].

1.2 Clinical diagnosis

Before the identification of *COH1* gene, Kivitie-Kallio and Norio observed 29 Finnish patients with an homogeneous Cohen syndrome phenotype and found the following features to be essential for the diagnosis: 1) nonprogressive psychomotor retardation, motor clumsiness and microcephaly; 2) characteristic facial features, including high-arched or wave-shaped eyelids, a short philtrum, thick hair, and low hairline (Fig. 1A); 3) childhood hypotonia and joint laxity (Fig. 1B); 4) progressive retinochoroidal dystrophy and myopia (Fig. 1D); 5) intermittent isolated neutropenia. Other findings were not essentials but they strongly supported the diagnosis. They included thick eyebrows and eyelashes, prominent root of the nose, prominent upper central incisors (Fig. 1A), enlarged corpus callosum, slender fingers (Fig. 1C) and cheerful disposition [4].

Subsequently, evaluation of *COH1* mutated patients from outside Finland revealed that the diagnostic criteria suggested by Kivitie-Kallio and Norio were important but not obligatory features [5,9]. Thus, Chandler et al. suggested the presence of at least two of the following major criteria in a child with significant learning difficulties: 1) facial gestalt, characterized by thick hair, eyebrows and eyelashes, wave-shaped, downward-slanting palpebral fissures, prominent, beak-shaped nose, short, upturned philtrum with grimacing expression on smiling (Fig. 1A); 2) pigmentary retinopathy (Fig. 1D); 3) neutropenia. Less specific but supportive criteria included early-onset progressive myopia, microcephaly, truncal obesity with slender extremities (Fig. 1B) and tapering fingers (Fig. 1C), and joint hyperextensibility [5].

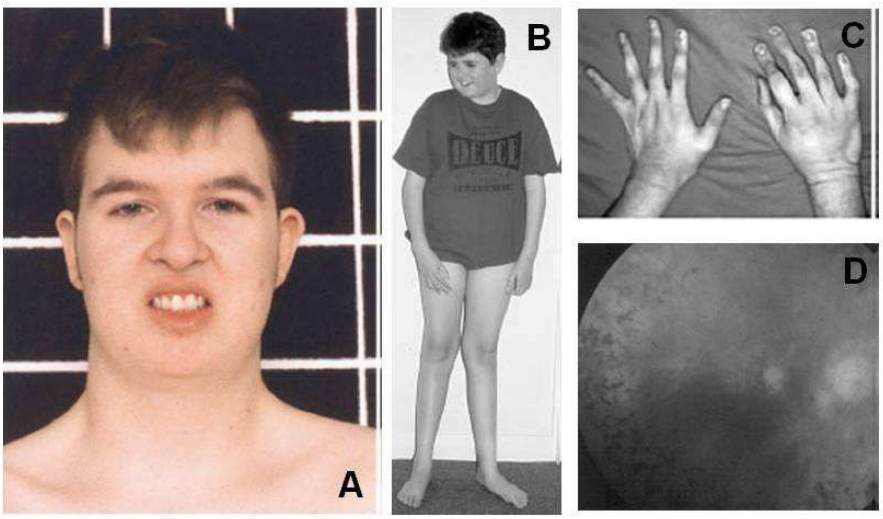

Figure 1. Phenotype of Cohen syndrome patients. **A**) Characteristic facial features include: downward-slanting palpebral fissures, thick hair and hairbrows, prominent nose with long columella, short philtrum and grimacing expression on smiling (from [8]); **B**) characteristic body with truncal obesity and slender extremities (from [5]); **C**) typical hands with tapering and slender fingers (from [6]); **D**) retinal dystrophy with waxy disc pallor and retinal pigmentation (from [6]).

As an aid to diagnosis, Kolehmainen et al. suggested the following 8 diagnostic criteria [6]:

- developmental delay
- microcephaly
- typical Cohen syndrome facial gestalt
- truncal obesity with slender extremities
- sociable behavior
- joint hypermobility
- high myopia and/or retinal dystrophy
- neutropenia

Patients fulfilling 6 or more criteria were considered likely to have Cohen syndrome, while those fulfilling ≤5 criteria were classified as “Cohen-like" patients. Using the mentioned criteria, Kolehmainen et al. found 22 different *COH1* mutations in probands diagnosed as "true Cohen syndrome" patients. By contrast, no *COH1* mutations were found in individuals who only met the provisional diagnosis of "Cohen-like syndrome” [6].

The broader clinical spectrum of the disease was confirmed by reports of Cohen syndrome patients from different ethnic groups with a wide geographical distribution, including British, German, Middle-Eastern, Amish, Japanese, Italian and Greek patients [7,10-14]. Mutational analysis in these patients demonstrated that *COH1* mutations may be present also in cases lacking some characteristic findings such microcephaly or truncal obesity [7,11,12]. Intellectual disability is an obligatory feature, but it varies from mild, moderate, to severe [15]. Progressive visual disability is a consistent finding, but retinopathy present at school age may not to be mandatory for the diagnosis of Cohen syndrome [7,12]. In contrast to the Finnish patients, where marked visual dysfunction occurred over 50 years of age, some British, German and Greek patients were described with total blindness [5,12,14]. Despite some craniofacial abnormalities including wave shaped palpebral fissures and short filtrum have been reported in almost all patients from different ethnic populations, the “facial gestalt” appears to be an unreliable indicator of Cohen syndrome and can lead to a delay in diagnosis [10].

Similar variation has been reported in affected people at different ages. In younger patients, the face is round with a full lower lip, the philtrum is not obviously short, the eyes are often slightly downward slanting with wave shaped eyelids, and the nasal bridge is not prominent (Fig. 2A). With increasing age, the facial appearance become more typical: the philtrum gets shorter, the longer columella further contributes to the impression of a short

filtrum, and the nasal bridge is higher [7,13] (Fig. 2B and 2C). Older patients maintain the typical features of Cohen syndrome, and thus can be used in making the diagnosis even at that age [12] (Fig. 2D).This high phenotypic variability leads to difficulty in establishing definitive clinical diagnostic criteria for Cohen syndrome patients.

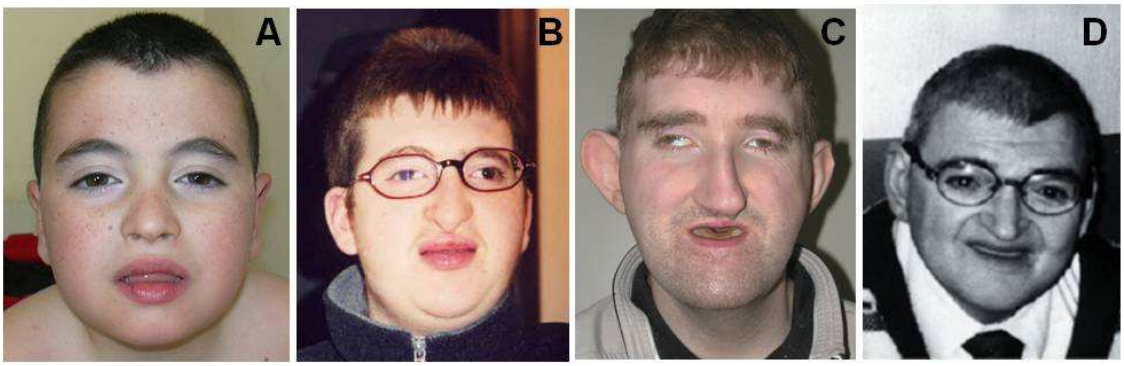

Figure 2. Typical facial features of Cohen syndrome patients at different ages. **A**) Five year old patient showing round face, down slanting eyes and thick eyebrows leading to a "China doll" appearance (from [13]); **B**) patient of 18 years showing shorter filtrum, longer columella and more prominent nasal bridge (from [13]); **C**) patient of 30 years: early loss of teeth contributes to maxillar hypoplasia (from [13]); **D**) patient of 44 years: typical facial gestalt is still present (from [12]).

1.3 *COH1* gene and encoded protein

In Finland, linkage analysis was performed in two-generation pedigrees with uniform clinical features, and the locus responsible of Cohen syndrome was mapped to the long arm of chromosome 8 [16]. In 2003, Kolehmainen et al. refined the critical region on 8q22 by haplotype analysis and characterized a new gene, *COH1*, that was mutated in Cohen syndrome patients [8].

By in silico analysis, exon prediction, and RT-PCR methods, Kolehmainen et al. obtained a full-length *COH1* cDNA. The longest transcript (14,093bp) is transcribed from

62 exons spanning a genomic region of 864 kb. The corresponding ORF encodes a protein of 4,022-amino acids and the translation initiation codon is at nucleotide 112 of the cDNA sequence, within exon 2 [8].

Overlapping EST sequences suggested that *COH1* has a complicated pattern of alternative splicing leading to the use of different termination codons and to additional in-frame, alternatively spliced forms [8]. There are two full-lenght transcripts comprising exons 1 to 62, including exon 28 (NM_017890) or 28b (NM_152564) (Fig. 3). Other less frequent variants have been described, originated by exon skipping, intron retention, or minor splicing changes resulting in alternative exons [17]. Most of them contain an early stop codon that leads to a truncated form (Fig. 3).

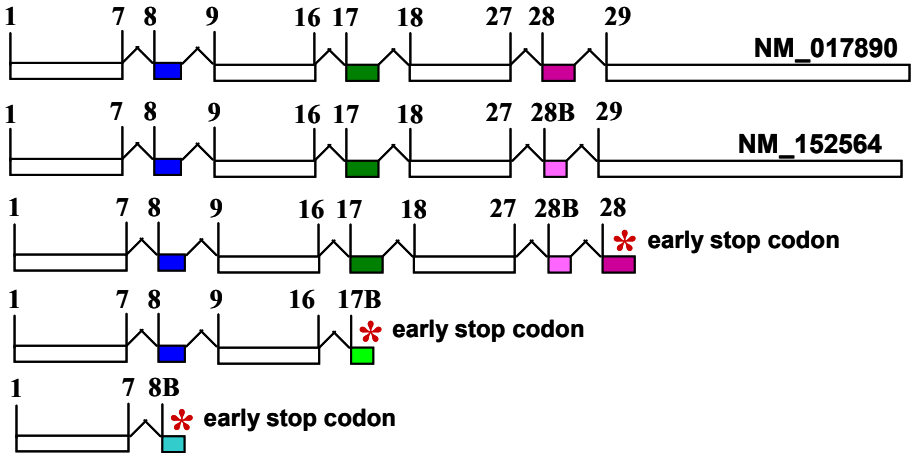

Figure 3. Schematic representation of *COH1* alternative splicing. Alternative use of exon 28 or 28b leads to the two full-length transcripts NM_017890 and NM_152564, differing for 25aa (4022aa and 3997aa respectively). Inclusion of 28 and 28b together, as well as the use of alternative exons 8b and 17b, lead to truncated splicing forms, since an early stop codon is introduced.

Northern blot analysis reveals a 14kb full-lengh transcript, demonstrating that both alternatively splice variants NM_017890 and NM_152564 encode functional proteins. We

can hypothesize that smaller transcripts, beeing only barely detectable, are either from imperfect splicing or play a regulatory role in *COH1* expression [15].

COH1 is widely expressed, with differential expression of different transcripts [8]. Variant containing exon 28b is the major, ubiquitously expressed *COH1* transcript, while variant containing exon 28 is predominant in brain [15,17]. The two main isoforms may have different functions within the cell or in different tissues. Since no *COH1* mutations affecting exons 28 or 28b have been identified, the relative importance of each splice variant has not been determined yet.

Expression analysis of *COH1* homologue in the mouse brain revealed that the gene is widely expressed in the central nervous system with no obvious regional differences, except for the glial cells where COH1 shows a lower expression. In addiction, it has been demonstrated that the level of *COH1* expression is higher in the postnatal brain than in the embryonic brain, suggesting that the gene primarily functions in postmitotic neurons and not in neuronal progenitor cells [11]. This findings are consistent with the fact that Cohen syndrome patients generally show normal head circumference at birth and develop microcephaly postnatally, probably by disrupting dendritic or axonal outgrowth.

N- and C-terminal sequences of *COH1* are conserved and share homology to *Saccharomyces cerevisiae* vacuolar sorting-associated protein VPS13 (Fig. 4), involved in trafficking of membrane proteins between the trans-Golgi network and the prevacuolar compartment. This suggests a role for COH1 protein in intracellular vesicle mediated sorting and transport of protein [8].

COH1 also shares homology to *CHAC*, the gene coding for the protein chorein and altered in patients with choreocanthocytosis (OMIM 200150) [18]. CHAC, COH1 and other two human proteins belong to a gene family homologous to yeast VPS13, which underwent

duplication events early in vertebrate evolution resulting in the four paralogues. Therefore, *COH1* gene has also been named *VPS13B* [17].

Recently, Seifert et al. provided the first molecular and functional characterization of the COH1 protein [19]. They demonstrated that *COH1* encodes a Golgi associated protein that co-localizes with the *cis*-Golgi marker protein GM130. In contrast to the previous prediction that COH1 was embedded into the membrane via 10 transmembrane helices [8], biochemical fractionation and partitioning experiments showed that COH1 is a peripheral membrane protein [19]. Localization to the Golgi is mediated by a C-terminal fragment of 315 aa, which comprises a hitherto uncharacterised Golgi targeting domain. COH1 depletion using RNAi causes fragmentation of the Golgi ribbon into mini-stacks, suggesting that the protein is required for maintenance of the Golgi morphology. Consistent with this, fibroblasts from Cohen syndrome patients carrying frameshift or nonsense mutations display a similar fragmentation of the Golgi complex, demonstrating that Golgi dysfunction contributes to Cohen syndrome pathology [19].

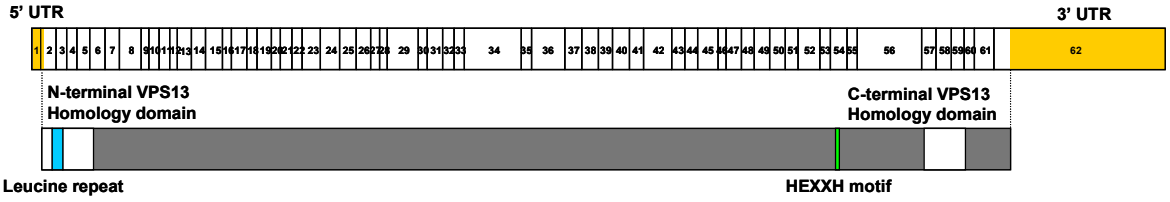

Figure 4. Schematic representation of *COH1* transcript (NM_017890) with the 62 exons and of the predicted protein [19]. 5' and 3' UTR are coloured in yellow. White boxes located at the N-terminal and C-terminal part of *COH1* indicates homolog regions to yeast VPS13. Other conserved patterns predicted by literature and Prosite searches are a Leucine repeat at the N-terminus (blue box) and a HEXXH motif at the C-terminus (green box).

1.4 *COH1* recurrent mutations

Mutations in *COH1* gene have been found in at least 200 patients from different ethnic backgrounds. Most of the alterations are nonsense or frameshift mutations leading to premature termination codons, while missense mutations are less frequent [12]. *COH1* intragenic deletions have also been reported [6,13,14,20]. Mutations occur throughout the gene and no mutational hotspots are observed. This extensive allelic heterogeneity explains the wide clinical variability in Cohen syndrome patients [7,12].

In Finland, where the phenotype is highly homogeneous, molecular analysis of *COH1* revealed a 2-bp deletion (c.3348_3349delCT) in exon 23 in most patients, resulting in a frameshift at codon 1117 and a stop that leads to protein truncation at codon 1124 [8]. Since Finnish population is structured in genetic isolates favouring inbreeding [7], this recurrent mutation is probably due to a founder effect.

In 2004, eight children with Cohen syndrome and homogeneous phenotype were reported in two extended Amish kindreds originating in northeast Ohio. Since marriages between consanguineous are very frequent in this population, the presence of a recurrent mutation due to a founder effect was hypothesised. In fact, all patients were homozygous for two different mutations in *COH1*: a 1-bp insertion in exon 51, resulting in a frameshift with a premature stop codon 19 amino acids downstream (p.L3087fsX19), and a 8459T>C transition in exon 46, resulting in a non-conservative substitution (p.I2820T) [10].

Cohen syndrome is also frequently observed among Irish travellers, a nomadic population that remained isolated for many years, practising consanguineous marriages. In this case, the recurrent allele carries the null mutation c.4471G>T (p.E1491X) in exon 29 of the *COH1* gene [21].

In Italy, a single base deletion in exon 58 (c.11125delC) has been reported, in heterozygous state, in three apparently unrelated patients coming from a restricted area of the Veneto's lowland between Padova town and Tagliamento river. This mutation leads to a stop codon in position 3769 of the protein (p.T3708fsX3769), cutting off the last transmembrane domain, likely giving rise to a non functional COH1 product. This could lead to think to a possible founder effect, but given the geographical conformation of this region, which is neither geographically or culturally isolated, a recent origin of the mutation could be hypothesized [13].

Recently, Bugiani et al. reported 14 Greek patients with a homozygous deletion spanning exon 6 to 16. The patients originated from two small neighbouring islands of the eastern Greek archipelago. Most of the affected individuals descend from consanguineous marriages among the offspring of a common Cretan ancestor. Even younger generations follow local tradition that privileges inbreeding, with a subsequent raise of Cohen syndrome incidence [14].

Most features associated so far with Cohen syndrome are not always present in patients sharing a founder mutation, thus representing clinical heterogeneity. Furthermore, there is a wide clinical variability among the different founder mutation cohorts, especially when evaluating separately specific system phenotypes, like the ophthalmic and central nervous systems [22].

1.5 Molecular analysis of *COH1*

Methods for the detection of point mutations in the *COH1* gene are well established in our laboratory and consist of denaturing high performance liquid chromatography (DHPLC) followed by automatic sequencing of the samples with abnormal elution profiles. In 2007, Katzaki et al. identified twelve *COH1* mutated alleles in 10 italian patients [13]. For the detection of possible large deletions/duplications not identifiable by sequencing analysis, seven Custom TaqMan Assays were designed for Real-Time quantitative PCR (qPCR) experiments. Using the probe for *COH1* exon 16, partial heterozygous gene deletions were identified in three patients from two different families. This was the first report in which gene deletions were investigated at genomic level [6,12,13].

However, as *COH1* is a large gene, spanning 846 kb of genomic DNA and composed by 62 exons, qPCR assays designed on a limited number of target regions are prone to miss a high fraction of intragenic rearrangements and do not allow the characterization of the extent of the deletions/duplications. In 2009, a targeted oligonucleotide array was designed, enabling the detection of *COH1* copy number changes with a median resolution of 200 bp. The authors analyzed 26 families with unexplained Cohen syndrome and identified deletions in 7 cases, showing that large deletions are an important cause of Cohen syndrome [20]. The authors proposed this technique as the method of choice for copy number screening of *COH1* gene in Cohen syndrome patients. However, since this method results quite expensive and not readily available, a new approach allowing the simultaneous analysis of different patients at a relatively low cost was needed.

2. Rationale and aim of the thesis

2. RATIONALE AND AIM OF THE THESIS

The present study focuses on the improvement of the molecular analysis of *COH1* gene in patients with a clinical diagnosis or suspect of Cohen syndrome. The cohort of patients collected by the Medical Genetics Unit of Siena from 2007 to 2011 include 96 cases. The collection of such a large cohort is due to the fact that the Unit of Siena is the unique centre performing Cohen syndrome molecular analysis in Italy. Moreover, given that there are very few Cohen syndrome reference centres in the world, we also receive DNA samples from different European countries such as France (6) and Spain (3).

Point mutation analysis was initially performed by DHPLC followed by automatic sequencing of cases with altered elution profiles. Using this method, 21 different point mutations were identified, including frameshift (6), splicing (3), missense mutations (11) and one complex rearrangement (Result 3.1). In order to evaluate the pathogeneticity of missense mutations not previously reported in literature, we considered the position of changed amino acids and conservation status of codons among different species (Result 3.1). Furthermore, we employed the prediction software Polyphen 2 (http://genetics.bwh.harvard.edu/pph2/) to evaluate the impact of the amino acid substitution on the structure and function of the protein (Result 3.1).

However, using these traditional techniques, a high percentage of patients showed only one or no mutated alleles. Considering the 96 patients and a group of 10 patients previously described by Katzaki et al. [13], we selected 14 cases with a phenotype strongly suggestive of Cohen syndrome, in which traditional tests failed to identify mutations in both alleles. Since in 2009 large rearrangements have been demonstrated to represent an important cause of Cohen syndrome by targeted array [20], we decided to set up a method

allowing a rapid and low cost *COH1* analysis of large deletions/duplications. We therefore chose to design, in collaboration with MRC-Holland (Amsterdam, The Netherlands, www.mlpa.com), a Multiplex Ligation-dependent Probe Amplification (MLPA) assay that could allow the screening and the characterization of *COH1* rearrangements. Using this method we identified deletions in 11 alleles and duplications in 4 alleles (Results 3.2). Considering our previous study [13], large rearrangements have been confirmed to represent a significant percentage of *COH1* mutations (42%). MLPA results were confirmed by Real Time quantitative PCR (Results 3.2). In order to better characterize the duplications and to predict the possible effect on the protein, in one case we performed long template PCR and automatic sequencing. This allowed to characterize the junction sequence of the duplicated segment and to predict that the complex rearrangement interrupts the protein product creating a premature stop codon (Results 3.2). Interestingly, we also found that 4 deletions showed the same exon coverage with respect to a deletion previously reported in a large Greek consanguineous family [14]. We therefore carried out haplotype analysis that suggested a possible founder effect in the Mediterranean basin (Result 3.2).

Furthermore, given that traditional *COH1* point mutation analysis is expensive and time-consuming, we decided to employ Next Generation Sequencing (NGS) on a Roche 454 GS Junior Sequencing platform to simultaneously screen the 62 exons of the gene (Result 3.3). We therefore set up the conditions for NGS analysis and we performed the mutation screening in the first Cohen syndrome patient (Results 3.3).

3. Results

3. RESULTS

3.1 Traditional point mutation analysis

3.1.1 Materials and Methods

Patients

The cohort of patients consisted in 96 cases with a clinical diagnosis or suspect of Cohen syndrome collected by the Medical Genetics Unit of Siena from 2007 to 2011. The cohort included: 84 patients from Italy, 6 from France, 3 from Spain, one from Holland, one from USA and one from Brazil.

Point mutation analysis

Blood samples were obtained after informed consent. Genomic DNA was isolated from peripheral blood using QIAamp DNA blood maxi kit, according to the manufactures' protocol (Quiagen, Hilden, Germany). All the 62 *COH1* exons were amplified by Polymerase Chain Reaction (PCR) in a reaction volume of 50 µl containing 100 ng of genomic DNA, 0.2 µM each primer, 0.2 mM dNTPs, 1X Buffer with 1.5mM MgCl2 and 1,5U/50µl Gold Taq polymerase (Applied Biosystems). Almost all primers used for *COH1* coding sequence amplification were already reported in literature (Kolehmainen et al, 2003), while they have been redesigned for exons 2, 8, 19, 22, 26, 29, 38, 47, 49, 56 e 57 (Table 1).

Amplicon	Forward primer (5' > 3')	Reverse primer (5' > 3')
2	----------	GATTACCGTCTAAACAAGCTG
8	----------	TCTACTGAATGCAAAGCAA
19	ATCAAATAAAGTTGAAATGTTATATTATG	GCAAAACAAGGGAATAATGATAG
22	CGTTTGGTATGTTCTGTG	----------
26	CATTTGCATGTAAGATGTGA	CAACAAGAGCAAAACTCTGT
29_2	ATCTCCTGGTCAGCCCATGA	ATATACCAAACCACAAAGCAC
38	GAACATAATTACAGTCCTAC	TGAAGAACTTCCCCTAAG
47	----------	CCCCAGTGCAAGGTTACTTT
49	----------	GATCATAAACGCAACTTTAC
56_1	----------	TTCCACGTATAACCGAGCA
56_2	CATATCCAACAAAGAGTTGG	----------
57	----------	AAGGAGTGAAGGCATTATTA

Tab. 1. New primers designed for *COH1* traditional point mutation analysis.

PCR conditions were as follows: 95 °C for 5 minutes, followed by 35 cycles consisting of melting at 95 °C for 1 minute, annealing for 1 minute at the primer-specific temperature (Table 2) and then extension at 72 °C for 1 minute, ended by a final elongation step at 72°C for 5 minutes. Amplicon size and purity were verified on a 1,2% agarose gel containing 0.5 μg/ml ethidium bromide.

Mutation analysis was performed by Denaturing High Performance Liquid Chromatography (DHPLC) using Transgenomic WAVE^{TM} apparatus (Transegenomic, San Jose, CA, USA). PCR products were denatured at 95°C, re-annealed at 65°C for 10 min and cooled at 4°C to generate heteroduplex. The optimal column temperature for fragments analysis was calculated using the WaveMaker Software (Transgenomic, San Jose, CA, USA) (Tab. 2). PCR products resulting in abnormal DHPLC profiles were purified and sequenced on both strands by using PE Big dye terminator cycle sequencing kit on an ABI Prism 310 genetic analyser (PE Applied Biosystems, Foster City, CA, USA).

Amplicon	Lenght (bp)	PCR Annealing temperature (°C)	DHPLC Column temperature (°C)	Amplicon	Length (bp)	Annealing temperature (°C)	Column temperature (°C)
1	451	56	64.7	32	329	52	54.4 /56.4
2	464	60	57.3	33	302	65	57.2
3	400	53	52.5	34-1	402	62	55.2 /56
4	425	59	55.3/58.6	34-2	560	62	55.3 /58.3
5	374	57	53.1/54.1	35	455	65	54.4 /55.4
6	315	65	53.8	36	599	65	55.3/57
7	344	54	55/53.7	37	545	67	54/ 56
8	732	53	52.2/53.4	38	794	60	51.1/53
9	321	57	51.6	39	563	60	53.4/55.2
10	342	58	54.9	40	387	65	56/57.2
11	271	57	54.9	41	393	54	56/58
12	271	58	54.6/55.6	42	603	52	53.7/54.7/ 57.6
13	314	65	53.7/57.2	43	468	59	58.8
14	301	58	50.4/53.4	44	495	62	55.1/56.1
15	278	57	55.6/58	45	549	62	54.9/55.9
16	255	58	55.7/59	46	337	58	54.5/57 (ts+1)
17	250	64	54.7/56.9 (ts+0.5)	47	541	59	55.6/56
18	462	57	54	48	418	65	53.5/55.5
19	320	58	52.3	49	518	59	53.3/55
20	522	59	52.5/54.5	50	399	56	52.6/56.9
21	453	65	57/57.4	51	510	60	52
22	269	54	52.6/53.6	52	508	60	56.4/57
23	547	58	51.7/54.7/57 (ts+1)	53	413	60	53.1/54
24	418	59	54.1/58.1 (ts+1)	54	380	62	56.2/59
25	410	63	53.3 e 56	55	354	65	54.6
26	731	55	52.1/56 (ts+1)	56-1	573	59	56/58
27	469	57	53.6/55.6	56-2	586	58	56.8/ 60.5 (ts +0.5)
28	659	61	53.5/54.5	57	309	57	58.8/60.8
28B	457	59	53.3	58	363	59	62.7/63.7
29_1	589	60	52.5/55	59	384	59	60.5
29_2	548	60	55	60	260	63	57.1/58.1
30	281	57	57.9	61	484	60	61/62.3
31	599	60	52/53	62	514	60	58

Tab. 2. Length of each amplicon of the *COH1* gene with relative annealing temperature for PCR amplification and optimal column temperature(s) for DHPLC analysis. ts: time-shift

In silico analysis

The free on-line software Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/) was used to predict the consequences of missense mutations [23]. Three empirically derived outcomes were used: probably damaging (it is with high confidence supposed to affect protein function or structure), possibly damaging (it is supposed to affect protein function or structure), and benign (most likely lacking any phenotypic effect). ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) was used to align COH1 amino acid sequences among different species in order to evaluate the conservation of mutated residues. Moreover, Splice Site prediction software of BDGP (Berkley Drosophila Genome Project, http://www.fruitfly.org/seq_tools/splice.html) was used to determine the predicted splice sites for both wild-type and mutated sequences.

3.1.2 Results

The molecular screening of the 96 patients by DHPLC revealed 21 different *COH1* sequence alterations in 22 patients (from 20 families).

Nine of these changes are surely pathogenetic mutations leading to a premature stop codon and resulting in a null allele. They include 6 frameshift and 3 splicing mutations (Tab. 3). All of them are never been reported before. In the two sibs C164 and C268, an in-frame rearrangement consisting of an insertion of two T and a deletion of 20 bases was identified on the paternal allele (Tab. 3). At protein level, seven amino acids are deleted and a Threonine is inserted.

Patient ID	Nucleotide change	Aminoacid change	Type of mutation	Inheritance
C91	- c.11695delAGTG	p.S3899fsX42	frameshift	*de novo*
C104	- c.11556insT - IVS24+2T4C	p.V3853fsX32 /	frameshift splicing	maternal paternal
C145	- c.402insT	p.L135fsX10	frameshift	maternal
C160/C161	- IVS4-2A>G	/	splicing	maternal
C167	- c.4474delA - IVS14-2A>G	p.I1492fsX42 /	frameshift splicing	paternal maternal
C164/C268	- c.5331insT - c.10880insTTdelCTGC GAGGCAGCTTGTGCAC	p.D1778X p.T3627_H3633 delinsT	frameshift in frame del-ins	maternal paternal
C185	- c.219delACinsT	p.K73fs93X	frameshift	maternal

Table 3. Frameshift and splice site mutations identified by traditional mutation analysis. The reference sequence of *COH1* gene is according to UCSC Genome Browser, http://genome.ucsc.edu, on Human March 2006 Assembly, hg18; NM_017890.

All the other alterations found are missense mutations (Tab. 4). Variations p.I1994V, present in patient C248, and p.S3142R, found in two unrelated patients C196 and C210, have already been reported by Seifert et al. as polymorphisms [12]. The missense change p.N2993S, identified in patient C121, has been already reported by Kolehmainen et al. [6]. Since this change creates an AG dinucleotide in the middle of exon 49, immediately upstream of a pyrimidine-rich region, a characteristic consensus of a splice acceptor site is created. Probably, an in-frame deletion of 37 aa in the conserved C-terminal domain of the protein is generated [6].

The other eight missense mutations have never been reported before (Tab. 4). Four (p.T1068I, p.T3375I, p.S3142G and p.T1894M) are non-conservative substitutions, changing the charge of the corresponding amino acid.

All the missense changes were analyzed by Polyphen-2 software. For variations p.N204S, p.N2993S, p.H520R and p.T1068I, the output was "probably damaging", while the others were predicted as "benign" (Tab. 4). In order to get more information about the evolutionary importance of each specific codon, we used ClustalW2 to align amino acid sequences of different species: for all the mutations predicted as "probably damaging", the corresponding residue results evolutionally conserved (Tab. 4); in the other cases, the position of the substitution is not or only partially conserved.

Patient ID	Nucleotide change	Aminoacid change	Inheritance	Reference	Polyphen output	ClustalW alignment
C114	- c.611A>G	p.N204S	maternal	/	probably damaging	conserved
C120	- c.8978A>G	p.N2993S	paternal	*Kolehmainen et al, 2004*	probably damaging	conserved
C121	- c.2704A>G	p.K902E	NA	/	benign	not conserved
C124	- c.1159A>G	p.H520R	NA	/	probably damaging	conserved
C137, C213, C260	- c.3203C>T	p.T1068I	paternal NA NA	/	probably damaging	conserved
C174	- c.10124C>T	p.T3375I	NA	/	benign	not conserved
C196, C210	- c.9424A>G	p.S3142R	NA	*Seifert et al, 2006*	benign	not conserved
C214	- c.5681C>T - c.7439G>C	p.T1894M p.C2480S	paternal maternal	/ /	benign benign	not conserved not conserved
C248	- c.5980A>G	p.I1994V	NA	*Seifert et al, 2006*	benign	not conserved
C297	- c.8672T>C	p.V2891A	NA	/	benign	not conserved

Table 4. **Missense variations identified by traditional point mutation analysis.**
blue: conservative changes; red: non-conservative changes; NA: not available data. The reference sequence of *COH1* gene is according to UCSC Genome Browser, http://genome.ucsc.edu, on Human March 2006 Assembly, hg18; NM_017890.

For each amplicon we determined the predicted splice sites by a Splice Site prediction software (BDGP). The comparison between wild-type sequences and mutated ones did not reveal any differences in splice sites.

Analysis of the parents revealed that all mutations are inherited, with the exception of p.S3899fsX42 in patient C91, that is a *de novo* alteration (Tab. 3).

3.1.3. Discussion

DHPLC technique allowed to identify 21 altered profiles in 96 patients addressed to *COH1* point mutation analysis. The subsequent analysis by automatic sequencing revealed 21 different variations, including frameshift, splicing and missense changes. The variations occurred throughout the gene, confirming the absence of mutational hotspot in *COH1* (Fig. 5).

Frameshift and splicing mutations are predicted to result in a premature stop codon generating a truncated protein product or causing degradation of *COH1* mRNA (nonsense-mediated RNA decay). For missense mutations, establishment of pathogeneticity has been more complex. In some cases, the nucleotide substitution causes the creation of a characteristic consensus sequence for a new acceptor or donor splice site, that could be used instead of the original one leading to a frameshift. This is the case for mutation p.N2993S, as predicted by Kolehmainen et al. [6].

A significant fraction of non-synonymous alterations located in coding regions is likely to affect protein structure or function [23]. Amino acid variants may interfere with folding, interaction sites, solubility or stability of the protein. These effects can be predicted from physical considerations and multiple sequence alignments [24].

Polyphen is an automatic tool for prediction of the impact of an amino acid substitution on the structure and function of a human protein [25]. This prediction is based on a series of features including the sequence, philogenetic and structural information characterizing the substitution. We decided to use the updated version of Polyphen, Polyphen-2, since it is demonstrated to be superior in predicting damaging mutations [26]. Four out of eleven mutations were estimated to be damaging with a high confidence. These changes correspond to residues strongly conserved across the species, indicating that the replacement may be incompatible with the spectrum of substitutions observed at the position in the family of homologous proteins. One of them, p.T1068I, also changes the residue from polar to non-polar, increasing the confidence for this substitution to have a damaging effect. On the other hand, given that this mutation is present in three unrelated and geographically distant patients (C137, C213, C260), there is the possibility that it represents a polymorphism.

Two mutations (p.S3142G and p.I1994V) have already been demonstrated to be polymorphisms, since they were detected in control samples [12]. According to software predictions, multiple alignments and rarity of the substitution, at least three (p.N204S, p.N2993S and p.H520R) of the other nine missense substitutions have good possibilities to be damaging. However, Polyphen-2 is a prediction software with a false positive rate of 20% [25] and could not lead to a definitive characterization of the substitution. In the absence of a functional assay, it remains possible that these missense changes represent rare nonpathogenic variants.

Since Cohen syndrome is a recessive condition, alterations in both alleles are necessary in order to confirm the clinical diagnosis. The employment of DHPLC followed by automatic sequencing has revealed the expected two alterations in only 4 cases (C104,

C164/C268, C167, C214). In 16 cases, only a single mutation was found. The other patients (76) included in this study did not show any alterations in the *COH1* gene. A small fraction of negative patients could be due to the limit of DHPLC technique that has a sensibility of 95-98%. Moreover, large rearrangements, not analyzed by traditional analysis, could account for negative patients. Alternatively, mutations could be present in the promoter or in other regulating sequences that are not included in the analysis. However, the high fraction of negative cases suggests a revaluation by clinical geneticists in order to decide whether the phenotype strongly indicate to continue with the molecular screening.

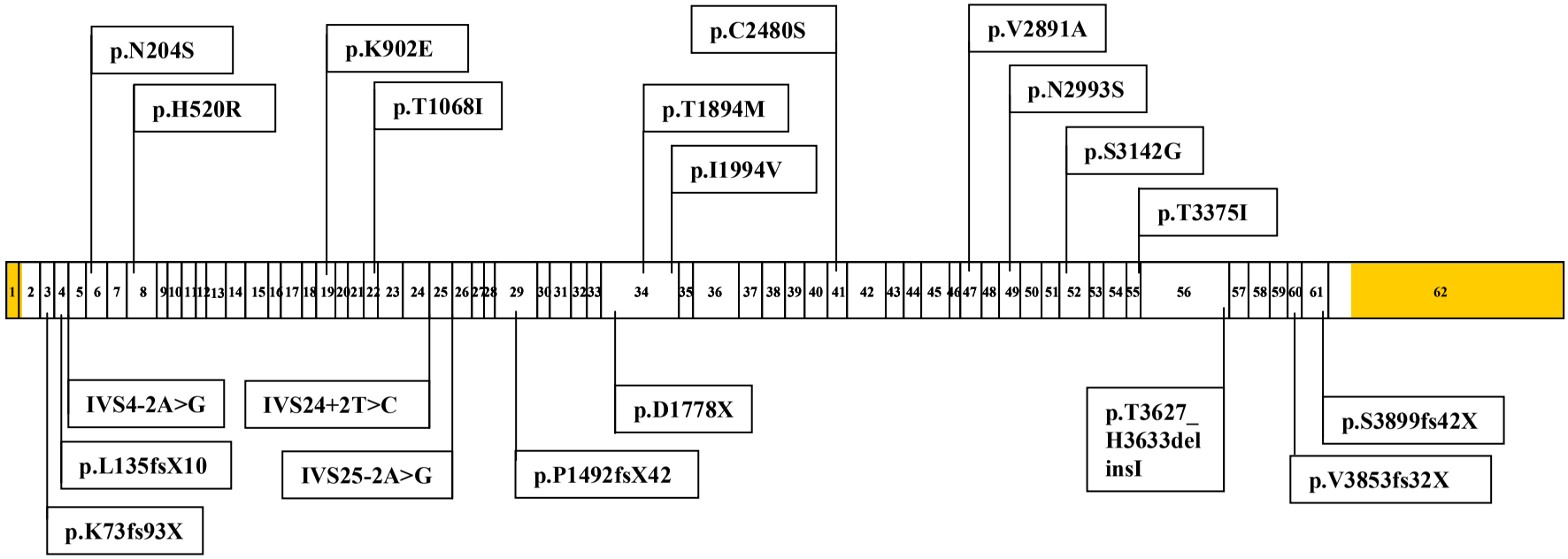

Fig. 5. Schematic representation of the *COH1* gene with distribution of the identified alterations. Relative length of each exon is represented. Yellow regions represent 5' and 3' UTRs. In the upper part, missense variations are reported, while frameshift and splicing mutations are indicated in the lower part.

3.2 High frequency of *COH1* intragenic deletions and duplications detected by MLPA in patients with Cohen syndrome

Veronica Parri, Eleni Katzaki, Vera Uliana, Francesca Scionti, Rossella Tita, Rosangela Artuso, Ilaria Longo, Renske Boschloo, Raymon Vijzelaar, Angelo Selicorni, Francesco Brancati, Bruno Dallapiccola, Leopoldo Zelante, Christian P Hamel, Pierre Sarda, Seema R Lalani, Rita Grasso, Sabrina Buoni, Joussef Hayek, Laurent Servais, Bert BA de Vries, Nelly Georgoudi, Sheena Nakou, Michael B Petersen, Francesca Mari, Alessandra Renieri, and Francesca Ariani

European Journal of Human Genetics (2010) 18, 1133–1140; doi:10.1038/ejhg.2010.59; published online 12 May 2010

European Journal of Human Genetics (2010) 18, 1133–1140

www.nature.com/ejhg

ARTICLE

High frequency of *COH1* intragenic deletions and duplications detected by MLPA in patients with Cohen syndrome

Veronica Parri[1], Eleni Katzaki[1], Vera Uliana[1], Francesca Scionti[1], Rossella Tita[1], Rosangela Artuso[1], Ilaria Longo[1], Renske Boschloo[2], Raymon Vijzelaar[2], Angelo Selicorni[3], Francesco Brancati[4], Bruno Dallapiccola[5], Leopoldo Zelante[6], Christian P Hamel[7], Pierre Sarda[7], Seema R Lalani[8], Rita Grasso[9], Sabrina Buoni[10], Joussef Hayek[10], Laurent Servais[11], Bert BA de Vries[12], Nelly Georgoudi[13], Sheena Nakou[14], Michael B Petersen[15], Francesca Mari[1], Alessandra Renieri*[,1] and Francesca Ariani[1]

Cohen syndrome is a rare, clinically variable autosomal recessive disorder characterized by mental retardation, postnatal microcephaly, facial dysmorphisms, ocular abnormalities and intermittent neutropenia. Mutations in the *COH1* gene have been found in patients from different ethnic origins. However, a high percentage of patients have only one or no mutated allele. To investigate whether *COH1* copy number changes account for missed mutations, we used multiplex ligation-dependent probe amplification (MLPA) to test a group of 14 patients with Cohen syndrome. This analysis has allowed us to identify multi-exonic deletions in 11 alleles and duplications in 4 alleles. Considering our previous study, *COH1* copy number variations represent 42% of total mutated alleles. To our knowledge, *COH1* intragenic duplications have never been reported in Cohen syndrome. The three duplications encompassed exons 4–13, 20–30 and 57–60, respectively. Interestingly, four deletions showed the same exon coverage (exons 6–16) with respect to a deletion recently reported in a large Greek consanguineous family. Haplotype analysis suggested a possible founder effect in the Mediterranean basin. The use of MLPA was therefore crucial in identifying mutated alleles undetected by traditional techniques and in defining the extent of the deletions/duplications. Given the high percentage of identified copy number variations, we suggest that this technique could be used as the initial screening method for molecular diagnosis of Cohen syndrome.
European Journal of Human Genetics (2010) **18**, 1133–1140; doi:10.1038/ejhg.2010.59; published online 12 May 2010

Keywords: Cohen syndrome; *COH1*; MLPA

INTRODUCTION

Cohen syndrome (OMIM 216550) is an autosomal recessive disorder first described in 1973 by Cohen *et al*.[1] It is characterized by non-progressive mental retardation, characteristic facial features, hypotonia, pigmentary retinopathy, myopia and intermittent neutropenia.[1–3] The peculiar craniofacial features of Cohen syndrome include microcephaly, downslanting and wave-shaped palpebral fissures, short philtrum and prominent upper central incisors.[1–3]

In 2003, mutations in the *COH1* gene were identified as causative of Cohen syndrome in the Finnish population.[4] The *COH1* gene maps to chromosome 8q22 and consists of 62 exons encoding for a potential transmembrane protein presumably involved in vesicle-mediated sorting and intracellular protein transport.[4,5]

The phenotypic spectrum in Finnish patients is highly homogeneous and molecular analysis revealed a founder effect with a common ancestral mutation causative of the majority of cases.[4] On the other hand, Cohen syndrome was found to be associated with mutations in the *COH1* gene in different populations with a broader clinical spectrum than the Finnish subtype.[4,6–10] About 100 mutations in the *COH1* gene have been identified so far.[9] Most of them are truncating mutations resulting in a null allele, whereas missense mutations and in-frame deletions are less frequent.[9]

Methods for the detection of point mutations in the *COH1* gene are well established in our laboratory and consist of denaturing high performance liquid chromatography (DHPLC) followed by automatic sequencing.[10] Until now, we used real-time quantitative PCR (qPCR) for the detection of large *COH1* deletions/duplications.[10] However, as *COH1* is a large gene, spanning 846 kb of genomic DNA and composed by 62 exons, qPCR assays designed on a limited number of target regions are prone to miss a high fraction of intragenic rearrangements and do not allow the characterization of the extent of the deletions/duplications. Very recently, a targeted oligonucleotide

[1]Medical Genetics, University of Siena, Siena, Italy; [2]MRC-Holland, Amsterdam, The Netherlands; [3]Paediatric Unit, University of Milan, Milan, Italy; [4]IRCCS Ospedale Casa Sollievo della Sofferenza, Laboratorio Mendel, San Giovanni Rotondo, Italy; [5]Ospedale Pediatrico Bambino Gesù, IRCCS, Rome, Italy; [6]Ospedale Casa Sollievo della Sofferenza, IRCCS, San Giovanni Rotondo, Italy; [7]Centre Hospitalier Régional et Universitaire, Service de Génétique Médicale, Montpellier, France; [8]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; [9]IRCSS MEDEA, Bosisio Parini, Lecco, Italy; [10]Child Neuropsychiatric Unit, University of Siena, Siena, Italy; [11]Institut de Myologie, Hôpital de la Pitié Salpétrière, Paris, France; [12]Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; [13]Department of Mental Health and Social Welfare, Institute of Child Health, Athens, Greece; [14]Department of Social and Developmental Pediatrics, Institute of Child Health, Athens, Greece; [15]Department of Genetics, Institute of Child Health, Athens, Greece
*Correspondence: Dr A Renieri, Medical Genetics, Department of Molecular Biology, University of Siena, V. Le Bracci 2, Siena 53100, Italy. Tel: +39 0577 233303; Fax: +39 0577 233325; E-mail renieri@unisi.it
Received 13 August 2009; revised 5 February 2010; accepted 25 February 2010; published online 12 May 2010

array was designed, enabling the detection of *COH1* copy number changes with higher resolution.[11] The authors analyzed 35 patients (from 26 families) with unexplained Cohen syndrome and identified deletions in 9 patients from 7 families, showing that large deletions are an important cause of Cohen syndrome.[11]

To detect *COH1* copy number variations, we used multiple ligation-dependent probe amplification (MLPA), a technique that has greatly improved mutation screening allowing the relative quantification of up to 40 different nucleic acid sequences in a single reaction tube at a relatively low cost.[12] By the use of two MLPA assays designed to screen copy number changes in almost all coding exons (60 out of 62) of *COH1*, we analyzed a group of patients with a clinical diagnosis of Cohen syndrome in which traditional tests failed to identify mutations in both alleles.

MATERIALS AND METHODS

Patients

Clinical geneticists from Italy, France, Holland and the United States assessed patients and diagnosed Cohen syndrome on the bases of published criteria.[13] Patients were considered as having Cohen syndrome when six of the following eight criteria were fulfilled: developmental delay, microcephaly, typical facial features, truncal obesity with slender extremities, sociable behavior, joint hypermobility, retinopathy or myopia, and intermittent neutropenia. Our series includes three children younger than 5 years (Table 1). As chorioretinal dystrophy does not manifest in young patients, the diagnosis of Cohen syndrome in children is considered when learning disabilities are associated with two of the following features: typical facial gestalt, pigmentary retinopathy or neutropenia.[14]

Overall, we collected 14 patients from 11 families, ranging in age from 18 months to 52 years. This group included four patients (1, 8, 9A, 9B) originally described by Katzaki *et al*[10] and 10 newly ascertained cases. The main clinical features are summarized in Table 1. Enrolled cases included one consanguineous family with an affected child (8) and 10 non-consanguineous families: 7 with one affected child (1, 2, 3, 4, 5, 6, 7), one with two affected sisters (9A, 9B), one with two affected brothers (10A, 10B) and one with an affected brother (11A) and sister (11B). A distinct phenotype was present in the two affected brothers (10A and 10B), presenting five of eight diagnostic criteria (Table 1); these patients were classified as Cohen-like.[13]

COH1 molecular analysis

Genomic DNA was isolated using QIAamp DNA blood maxi kit, according to the manufacturer's protocol (Qiagen, Hilden, Germany). PCR amplification of the 62 exons was carried out using published primers.[4,10] Mutation analysis was performed by DHPLC using the Transgenomic WAVE (Transgenomic, San Jose, CA, USA).[10] Quantitative PCR was also performed in one familiar case (9A, 9B) and one sporadic case (8) with a Custom TaqMan Assay designed on exon 16 (Applied Biosystems, Foster City, CA, USA).[10]

MLPA analysis was performed using two distinct SALSA MLPA kits (P321-A1/P322-A1) designed by MRC-Holland (Amsterdam, The Netherlands). The two assays include 69 *COH1* probes to screen copy number changes in almost all coding gene exons (60 out of 62) and 16 control probes. No probe was present for exons 6 and 14. For exons 3, 16, 17, 24, 31, 34, 35 and 36, two distinct probes were designed. The analysis was carried out as previously described.[12] Briefly, 100 ng of genomic DNA was diluted with TE buffer to 5 μl, denatured at 98°C for 5 min and hybridized with SALSA Probe-mix at 60°C overnight. Ligase-65 mix was then added and ligation was performed at 54°C for 15 min. The ligase was successively inactivated by heating the samples at 98°C for 5 min. PCR reaction was performed in a 50 μl volume. Primers, dNTPs and polymerase were added and amplification was carried out for 35 cycles (30 s at 95°C, 30 s at 60°C and 60 s at 72°C). The amplification products were separated on an ABI Prism 310 automatic sequencer and analyzed using the GenScan software ver.3.1 (Applied Biosystems). For data analysis, the values of peak sizes and areas were exported to an Excel table and compared with a normal control (MRC-Holland). Dosage alterations were considered significant if sample values deviated more than 30% from the control.

For exons 6 and 14, we designed two specific qPCR assays (Supplementary Table 1). In addition, MLPA results were confirmed by qPCR using probes located in exons 16, 24, 34, 42, 48 and 58 (Supplementary Table 1).[10] Reactions were performed in a 96-well optical plate with a final reaction volume of 50 μl using an ABI prism 7000 (Applied Biosystems). A total of 100 ng of DNA (10 μl) was dispensed in each of the four sample wells for quadruplicate reactions. Thermal cycling conditions included a pre-run of 2 min at 50°C and 10 min at 95°C. Cycle conditions were 40 cycles at 95°C for 15 s and 60°C

Table 1 Summary of the clinical features in Cohen patients

Case	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9A*	*9B*	*10A*	*10B*	*11A*	*11B*
Patients ID	C8[a]	C91	C104	C145	C155	C167	C185	R111[a]	C42[a]	C43[a]	C160	C161	C164	C268
Sex	M	F	F	F	F	F	F	M	F	F	M	M	M	F
Consanguineous parents	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No
Age at assessment	5 years	20 years	10 years 6 months	19 years	17 years	19 years	3 years 6 months	6 years 3 months	52 years	51 years	45 years	40 years	4 years 6 months	2 years 4 months
Mental retardation (degree)	Yes	Severe	Mild-moderate	Moderate	Moderate	Mild-moderate	Moderate	Moderate	Yes	Yes	Moderate	Moderate	Moderate	Moderate
Microcephaly	+	+	+	+	3° cnt	+	+	+	+	+	–	–	+	+
Typical facial gestalt	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Truncal obesity	–	–	+	+	+	+	–	+	+	+	–	–	–	–
Narrow H/F; slender/ tapering fingers	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Retinopathy	+	+	+	+	+	+	+	+	+	+	+	+	–	–
Myopia (diaptres)	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Neutropenia	+	+	-	+	+	+	+	+	+	+	+	+	+	+
Joints hyperlaxity	+	+	+	+	+	+	+	-	+	+	+	+	+	+
Sociable behavior	–	NR	+	+	NR	+	+	+	+	+	NR	NR	+	+
Other	Pes varus	Mild mitral insufficiency	Leg asymmetry	IUGR hip asymmetry			Neonatal hypotonia	Syndactyly (II-III toes)	Breast cancer, bilateral cataract	Breast cancer, bilateral cataract	Mitralic insuffi-ciency			

NR, not reported.
[a]Patients already reported.[10]

for 1 min, according to the TaqMan Universal PCR Protocol (PE Applied Biosystems). The TaqMan Universal PCR Master Mix and Microamp reaction tubes were supplied by Applied Biosystems. The starting copy number of the unknown samples was determined using the comparative C_t method, as previously described.[13]

In case 11, long-range PCR was performed with the Expand Long Template PCR kit as specified by the manufacturer (Roche Diagnostics, Basel, Switzerland), using a forward primer located in intron 59 (ggatggctctgaacagatga) and a reverse primer located in intron 56 (agaagcaattggcaagaggt). These primers are divergent in the normal genome and they do not amplify the control's DNA. PCR conditions were as follows: 300 nM of each primer, 350 μM of dNTPs, 2.0 mM $MgCl_2$, 0.75 μl of enzyme mix and 1× Buffer II, and the following cycling parameters: 94°C 5 min; 94°C 10 s, 59°C 30 s, 68°C 5 min, 10 cycles; 94°C 15 s, 59°C 30 s, 68°C 5 min +20 s/cycle, 25 cycles; final extension 68°C 30 min.

Haplotype analysis

A set of 10 markers covering a region of about 4 Mb encompassing the *COH1* gene were used for haplotype analysis (Supplementary Table 2) in three of our cases with the 6–16 deletion (case 5, 8 and 9A) and one member of the large Greek consanguineous family reported by Bugiani *et al*[16] harboring the 6–16 deletion in homozygous state. Haplotype analysis was also performed in all available family members of the 6–16 deleted patients and in 50 Italian control individuals. The forward primers were fluorescently labeled with FAM. Markers were amplified by polymerase chain reaction. Conditions were optimized for individual primer pairs in a 9600 thermocycler (Applied Biosystems). The programs used were 95°C for 12 min, followed by 30 cycles of melting at 94°C for 15 s, annealing at the optimal temperature for 15 s, and then extension at 72°C for 30 s. A final extension was performed at 72°C for 10 min. PCR products were run on an ABI 3130 sequencer (Applied Biosystems) and analyzed with GeneMapper v.4.0. The size of the PCR products of the microsatellite markers were compared among the families carrying the recurrent deletion 6–16 in heterozygous or homozygous state, in order to define the haplotype co-segregating with the deletion.

RESULTS

Phenotype

All 14 patients displayed the typical Cohen facial gestalt, narrow extremities and truncal adiposity even if not all cases were obese (7 out of 14) (Table 1, Figure 1).[10] Microcephaly was present in the majority of patients (9 out of 14) (Table 1).[10] The retinopathy was absent in one family with two affected children younger than 5 years (11A, 11B) (Table 1). Neutropenia was absent in one patient (3) and one case did not show joint hyperextensibility (8). Among the 14 patients, two brothers (10A, 10B) presented an atypical phenotype, lacking microcephaly and truncal obesity. However, the diagnosis of Cohen syndrome was suggested based on the association of retinopathy, neutropenia and facial appearance (Figure 1).

COH1 molecular analysis

The 14 patients (11 families) with a clinical diagnosis of Cohen syndrome were analyzed for the presence of *COH1* point mutations by DHPLC followed by sequencing of the samples with an abnormal elution profile.[10] This analysis led to the detection of 12 different mutations, including six frame-shift, three splice site, two nonsense and one complex rearrangement (Table 2). Moreover, in one family (9A and 9B) and in one sporadic patient (8) a partial heterozygous *COH1* gene deletion was already detected by qPCR using a TaqMan probe designed on exon 16.[10] To identify missed mutated alleles and to characterize the extent of the deletions/duplications, we used two MLPA assays (P321-A1/P322-A1) designed to detect *COH1* copy number changes in 60 out of 62 exons of the gene. This method led us to identify 5 different multi-exonic deletions in 11 alleles and 3 different duplications in 4 alleles (Table 2). In particular, MLPA characterized heterozygous copy number variations in nine patients (seven families) displaying a point mutation previously identified by

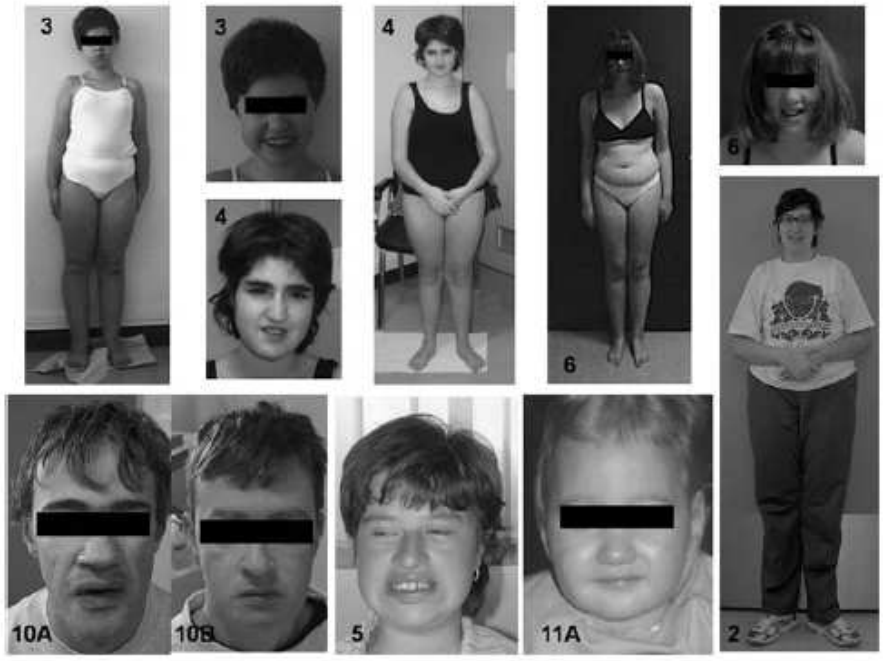

Figure 1 Clinical features of Cohen syndrome patients. Note the typical facial gestalt of patients 3, 4, 5, 6, 10A, 10B and 11A. Frontal views of patients 2, 3, 4 and 6, showing truncal obesity.

Table 2 *COH1* point mutations and large deletions/duplications identified in the study

Case	*Patient ID*	*Nucleotide change*	*Amino-acid change*	*Copy number change*	*Inheritance*
1	C8[a]	c.3427C>T	p.R1143X	DelEX32-35	M
					P
2	C91	c.11695delAGTG	p.S3899fsX42	DupEX4-13	*De novo*
					P
3	C104	c.11556insT;	p.V3853fsX32		M
		IVS24+2T>C			P
4	C145	c.402insT	p.L135fsX10	DupEX20-30	M
					P
5	C155			DelEX6-16[b]	NA
6	C167	c.4474delA;	p.I1492fsX42		P
		IVS14-2^{A}>G			M
7	C185	c.219_20delACinsT	p.K73fsX20	DelEX40-43	M
					P
8	R111[a]	c.11564delA	p.Y3855fsX22	DelEX6-16	P
					M
9A/B	C42/C43			DelEX6-16	NA
				DelEX46-50	NA
10A/B	C160/161	IVS4-2A>G		DelEX4-16	M
					P
11A/B	C164/268	c.5331insT;	p.D1778X	DupEX57-60	M
		c.10880insTTdelCTGCGA	p.T3627_H3633delinsI		P
		GGCAGCTTGTGCAC			

NA, not available for testing; P, paternal; M, maternal.
[a]Patients previously described.[16]
[b]In homozygous state (The reference sequence of *COH1* gene is according to UCSC Genome Browser, http://genome.ucsc.edu, on Human March 2006 Assembly, hg18; NM_017890).

DHPLC on the other allele (1, 2, 4, 7, 8, 10A, 10B, 11A and 11B), two different compound heterozygous deletions in two affected sisters (9A and 9B) and one homozygous deletion in one sporadic patient (case 5) (Table 2, Figure 2, Supplementary Figures 1 and 2).

In four patients, MLPA showed the presence of a deletion spanning from exons 7 to 16 (Figure 2, Table 2). As the MLPA assays contain 69 probes not including exon 6, we designed a targeted qPCR probe assay for this exon (Supplementary Table 1). This analysis showed that the four deletions spanned indeed from exons 6 to 16 (Figure 3, Table 2).

In two sporadic patients (cases 2 and 4), MLPA detected a significant increase in the fluorescent signals corresponding to exons 4–13 and 20–30, indicating the presence of two differently sized duplications (Table 2, Supplementary Figure 1). In case 2, a specific qPCR assay indicated that exon 14 is not included in the duplication (data not shown). In a familial case (11A, 11B) in which DHPLC followed by sequencing had already detected a complex rearrangement in exon 56 (c.1088insTTdelCTGCGAGGCAGCTT GTGCAC; p.T3627_H3633delinsI), MLPA also disclosed a significant increase in peak heights 57–60, suggesting the presence of a heterozygous duplication (Table 2, Figure 4). Analysis of the parental DNA indicated that the rearrangement p.T3627_H3633delinsI was in *cis* with the duplication detected by MLPA (Table 2).

To better characterize the 57–60 duplication, we performed long-range PCR using a forward primer in intron 59 and a reverse primer in intron 56 (Figure 5a). We obtained a product of ~1 kb in the two affected sibs and in the carrier father. Automatic sequencing of the PCR product permitted us to characterize the junction sequence of the duplicated segment (Figure 5), 95 bp downstream with respect to the rearrangement. The duplicated segment, starting within intron 56, is inserted within exon 61 in position g.100953994 (NM_017890) (Figure 5). According to prediction software, this insertion interrupts the protein product creating a premature stop codon after 10 new amino acids.

Not all parental DNAs were available for testing (Table 2). For patients 9A and 9B, the DNAs of two healthy sibs have been analyzed to determine whether the two rearrangements were in *cis* or *trans*. MLPA revealed that the brother and the sister were carriers of the deletions spanning exons 6–16 and exons 46–50, respectively, confirming that the rearrangements were in compound heterozygosity. In the cases where parental DNAs have been tested, all mutations were inherited except in one patient (case 2) harboring a *de novo* point mutation (c.11695delAGTG; p.S3899fsX42) (Table 2).

All copy number changes identified by MLPA were confirmed by qPCR using specific probes for exons 16, 24, 34, 42, 48 and 58 (data not shown).

Haplotype analysis

To investigate a founder effect for the recurrent deletion of exons 6–16, we performed haplotype analysis in three of our cases and one additional case belonging to a large Greek consanguineous family reported by Bugiani *et al*.[16] A founder effect is expected to result in sharing of allelic sequence polymorphisms in the vicinity of the deletion. We examined 10 microsatellite markers within a region of about 4 Mb encompassing the *COH1* gene (Supplementary Table 2, Table 3). For heterozygous markers, the phase was assigned by genotyping other family members: parents in case 8 (a carrier mother and noncarrier father) and sibs in case 9 (one carrier and one noncarrier sister) (data not shown).

To determine how frequently alleles of the same size can be obtained by chance in a general population, we genotyped DNA from 50 Italian control samples using primers for the same 8 microsatellite markers

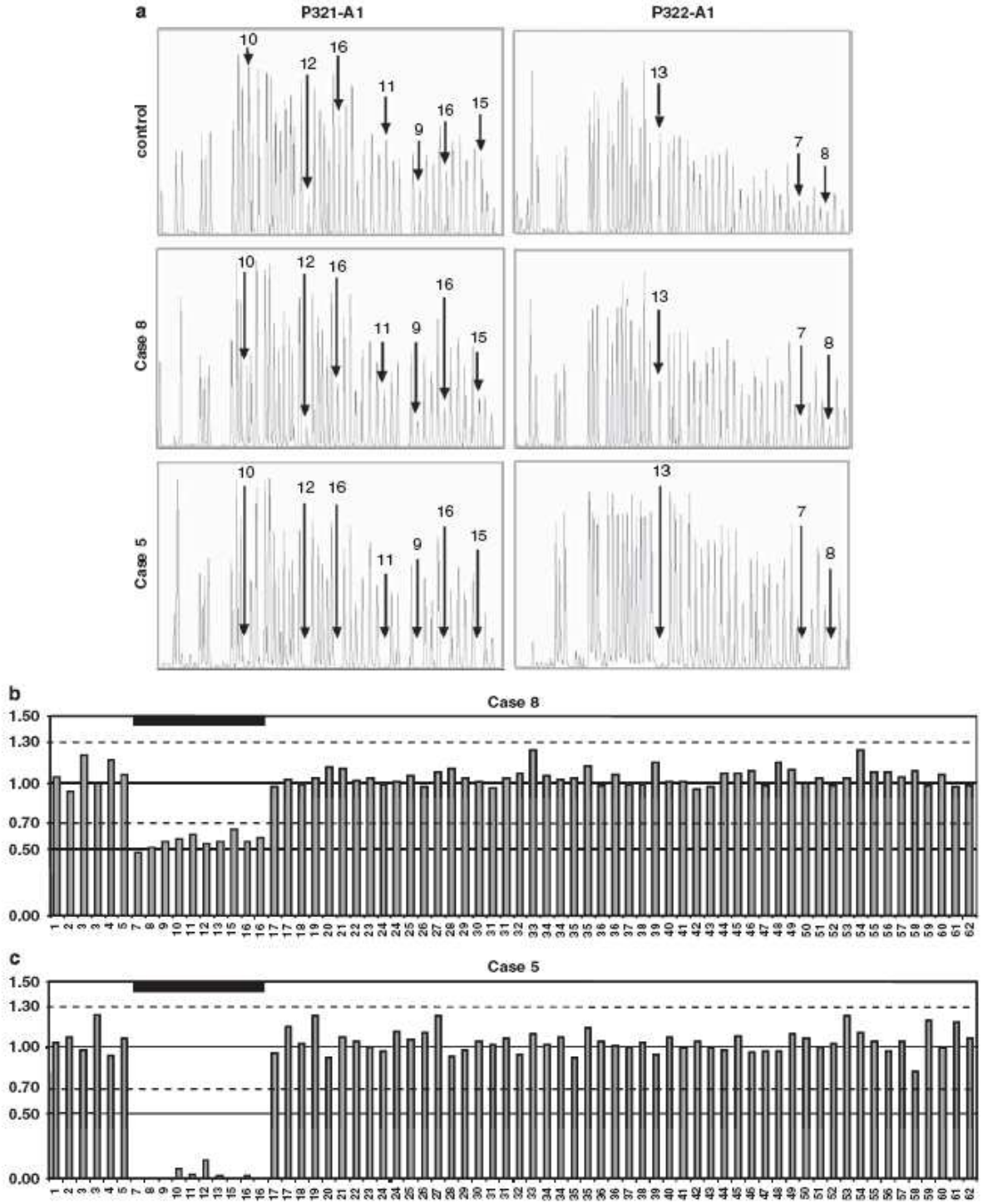

Figure 2 MLPA analysis results showing the recurrent deletion in heterozygous (Case 8) and homozygous (Case 5) states. (a) Electropherograms obtained with P321-A1 kit (on the left) and P322-A1 kit (on the right) for a normal control sample, patient 8 and patient 5. Numbers and arrows indicate the exon probes with reduced fluorescence signals with respect to the control sample. In patient 8, the signal is half-reduced for probes 7–16, whereas in patient 5 there is no signal for the same probes. (b, c) Peak area histograms for patients 8 (a) and 5 (b) normalized with the control sample. Exon dosage is reported on the *y* axis (normal values spanning from 0.7 to 1.3 are indicated with broken lines). MLPA analysis shows reduced peak area for exons 7–16, compatible with a heterozygous deletion in patient 8 and a homozygous deletion in patient 5. Deletions are indicated with a heavy black line.

(minimal common haplotype, Table 3). None of the healthy controls and none of the noncarrier family members showed the minimal common haplotype (data not shown).

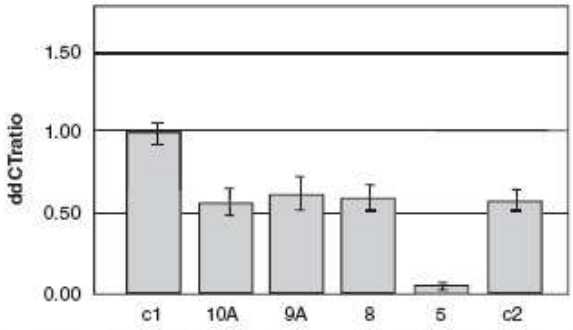

Figure 3 Quantitative PCR results for exon 6 of *COH1*. ddCT ratios and standard deviations of a normal control sample (c1), a deleted control sample (c2) and patients 5, 8, 9A and 10A. Compared with controls, patients 8, 9A and 10A show ddCT ratio values of about 0.5, indicating a deletion in the heterozygous state, whereas patient 5 shows ddCT ratio values of about 0.0, indicating a deletion in the homozygous state.

DISCUSSION

In this study, we report the first application of the MLPA technique to screen for *COH1* large deletions and duplications. In a group of 14 patients (11 families) with a clinical diagnosis of Cohen syndrome, MLPA allowed us to obtain rapid and high quality results disclosing 11 deleted and 4 duplicated *COH1* alleles. The use of MLPA led us to identify all *COH1* mutations undetected by conventional screening, suggesting that this technique is an important tool for the molecular characterization of Cohen syndrome.

Our series included 12 patients with true Cohen syndrome and two brothers with an atypical phenotype, lacking microcephaly and truncal obesity. However, the association of retinopathy, neutropenia and facial appearance addressed the clinical diagnosis. Their facial features, although not typical, were not in disagreement with the diagnosis of Cohen syndrome consisting of long face, heavy eyebrows, mildly down-slanting palpebral fissures, prominent root of the nose, normal philtrum and prognatism (Figure 1). Three patients from two families were children aged less than 5 years. They presented the typical facial features of younger patients, including round face with full lower lip, not excessively short philtrum, slightly downward-slanting eyes with wave-shaped eyelids and less prominent nasal bridge (Figure 1).[10]

Copy number changes in *COH1* have been previously investigated in patients with Cohen syndrome by qPCR using probes designed on a

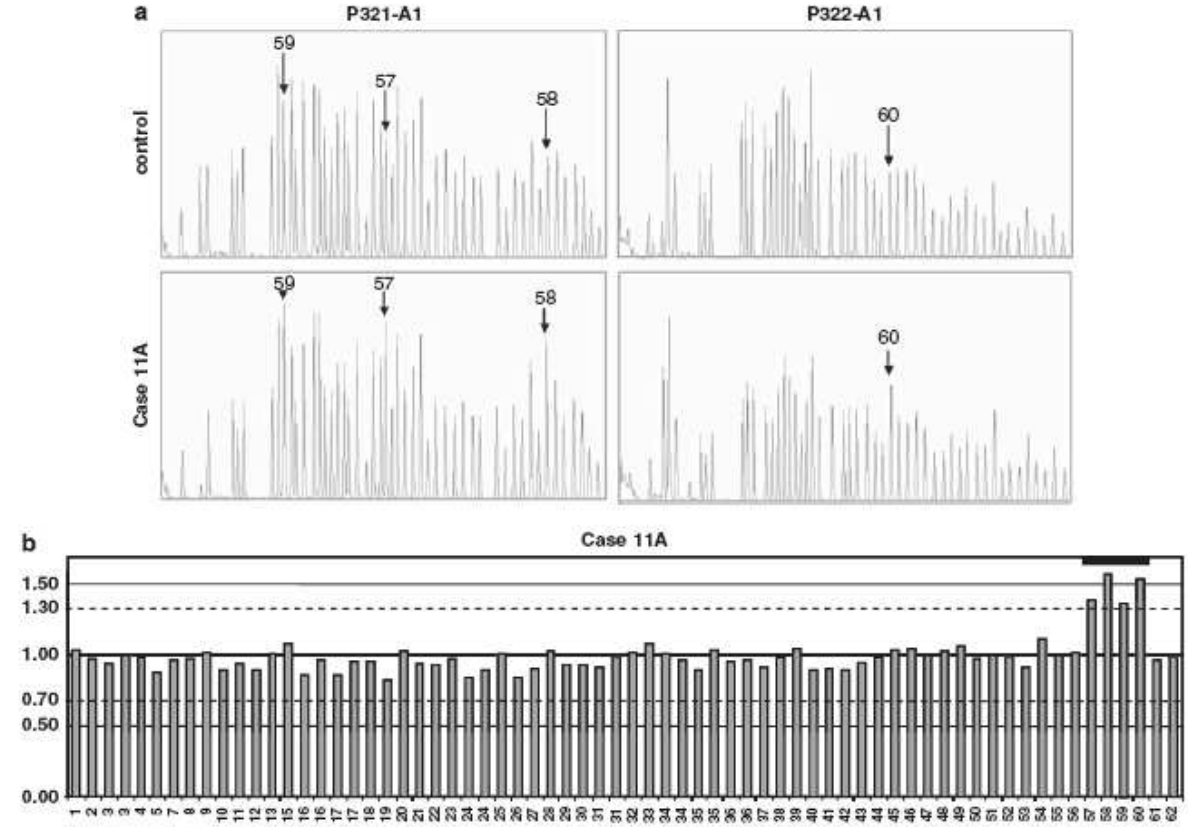

Figure 4 MLPA analysis results showing the duplication spanning exons 57–60 in the familial case with an affected brother (11A) and sister (11B). (a) Electropherograms obtained with P321-A1 kit (on the left) and P322-A1 kit (on the right) for a normal control sample and patient 11A. Numbers and arrows indicate the exon probes with increased fluorescence signals with respect to the control sample. (b) Peak area histograms for patient 11A normalized with the control sample. The exon dosage is reported on the *y* axis (normal values spanning from 0.7 to 1.3 are indicated with broken lines). The consistent increase in the peak area for exons 57–60 is compatible with a duplication of these exons (indicated with a heavy black line).

limited number of exons.[10,16] Only recently, a targeted oligonucleotide array with a median resolution of 200 bp was designed within the gene, which considerably increased the mutation detection rate.[11] Using this technique, the authors identified *COH1* large deletions in nine patients from seven families, showing that they represent an important cause of Cohen syndrome.[11] The present results and our previous study on a group of 18 patients disclosed a total of 21 alleles with point mutations (58%) and 15 with copy number variations (42%), confirming that deletions and duplications account for a significant percentage of *COH1* mutations.[10]

In four patients from three families, MLPA identified a *COH1* large deletion sharing the same extent with one previously reported in an isolated Greek Island population, spanning from exons 6 to 16.[16] In our patients, the deletion was heterozygous in two families and homozygous in an apparently non-consanguineous family.[10] Interestingly, this latter patient displays the same constellation of facial features reported in Greek patients with homozygous deletion including thick hair with low hairline, strabism, lack of nasofrontal angle, short upturned philtrum and prominent maxillary central incisors (patient 5, Figure 1).[16] Moreover, they show milder microcephaly and more severe visual impairment than the original phenotype described in the Finnish population.[4,16]

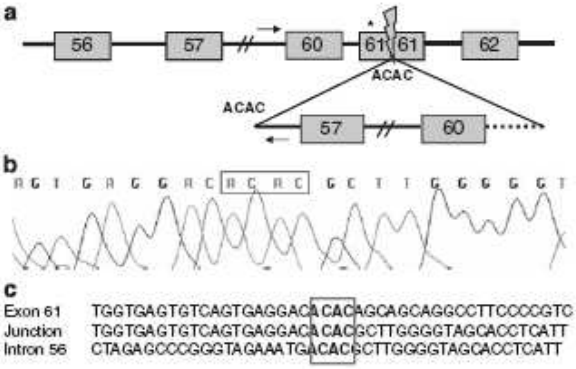

Figure 5 Characterization of duplication 57–60 in familial case 11. (a) Schematic drawing of the duplicated region. The star indicates the position of the MLPA probe in exon 61, whereas the thunder represents the insertion point of the duplicated segment. Arrows indicate the primers located within introns 59 and 56 used in the long-range PCR experiment. (b) Sequence analysis showing the junction between intron 56 and exon 61. (c) Aligned exon 61 and intron 56 sequences at the duplication junction. Region of homology across the duplication junction is boxed.

Our three families with the same deletion encompassing exons 6–16 come from different Italian regions, two in Central Italy and one in Southern Italy. The results obtained by haplotype analysis in these families, in one member of the large Greek consanguineous family previously reported by Bugiani *et al*[16] harboring the 6–16 deletion in homozygous state and in 50 healthy Italian controls, suggest that the recurrent deletion is due to an ancestral founder effect in the Mediterranean area (Table 3).

In this study, we also identified two deletions spanning exons 4–16 and 40–43, sharing the same exon coverage with two deletions already reported in the Northern European population.[11] Also, in these cases we cannot exclude a founder effect for the deleted alleles. Alternatively, these could be independent mutations favored by the presence of repeated elements located at the break points. Accordingly, RepeatMasker software analysis of the genomic region containing *COH1* revealed a higher frequency of LINEs, SINEs and DNA repeat elements in comparison with the average for autosomal sequences.[11] In a previous study, it was suggested that the most likely mechanism for genomic rearrangements in the *COH1* gene is the non-homologous end joining, leading to non-recurrent deletions.[11] Considering our latest results, the non-allelic homologous recombination mechanism cannot be ruled out.

In four patients from three families, MLPA identified three different size duplications spanning exons 4–13, 20–30 and 57–60, respectively. To our knowledge, *COH1* intragenic duplications have never been reported in Cohen syndrome.

In one family with two affected sibs (cases 11A/B), we identified a complex rearrangement (p.T3627_H3633delinsI) in *cis* with the downstream duplication detected by MLPA. We initially hypothesized that this rearrangement could be located at the break point of the duplication within exon 56. However, sequencing analysis of the long PCR product using a forward primer in intron 59 and a reverse primer in intron 56 indicated that the duplication effectively starts in intron 56, 95 bp after the rearrangement (Figure 5). This sequence is joined to exon 61 in position g.100953994 (NM_152564) (Figure 5). As the MLPA probe of exon 61 is located upstream of the junction point (Figure 5a) and its signal does not increase, we can suppose that the duplication is not in tandem. The insertion of the duplicated segments within exon 61 creates a premature stop codon after 10 new amino acids of the protein product. Even if detailed mapping of the extent of all the duplications has not yet been undertaken, these rearrangements

Table 3 Haplotype analysis in patients harboring the recurrent exons 6–16 deletion

Marker	*Position (Mb)*	*C37 (Greek)*		*Case 5 (Italian)*		*Case 9A (Italian)*		*Case 8 (Italian)*	
D8S1018	97598	315	319	315	315	319	323	319	315
D8S257	99451	109	109	109	109	109	—	109	113
8-23TC	99924	214	214	214	214	214	218	214	204
8-25GT	100056	353	353	353	353	353	379	353	351
8-20TG	100601	169	169	169	169	169	173	169	173
VPS13B	—	*del6_16*	*del6_16*	*del6_16*	*del6_16*	*del6_16*	*del46_49*	*del6_16*	*Y3855fsX22*
D8S1789[a]	100738	255	255	255	255	255	255	255	255
D8S470[a]	100743	226	226	226	226	226	226	226	226
D8S300	100987	485	485	485	485	485	499	485	499
8-18AC	101066	95	95	95	95	95	97	95	97
D8S398	101588	141	141	141	141	137	141	137	141

[a]Intragenic markers.
Grey columns: haplotype co-segregating with the deletion.
(The reference sequence of *COH1* gene is according to UCSC Genome Browser, http://genome.ucsc.edu, on Human March 2006 Assembly, hg18; NM_017890).

probably led to a frameshift and a premature truncation of the protein at different levels.

In conclusion, our study confirms that *COH1* copy number variations are a frequent cause of Cohen syndrome and consist of intragenic deletions as well as duplications. Therefore, incorporation of detection tools for *COH1* copy number variations is mandatory in the molecular diagnosis of Cohen syndrome.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank all the families for their participation in this study. We also thank MRC-Holland for the technical support.

1 Cohen Jr MM, Hall BD, Smith DW, Graham CB, Lampert KJ: A new syndrome with hypotonia, obesity, mental deficiency, and facial, oral, ocular, and limb anomalies. *J Pediatr* 1973; 83: 280–284.
2 Carey JC, Hall BD: Confirmation of the Cohen syndrome. *J Pediatr* 1978; 93: 239–244.
3 Norio R, Raitta C, Lindahl E: Further delineation of the Cohen syndrome; report on chorioretinal dystrophy, leukopenia and consanguinity. *Clin Genet* 1984; 25: 1–14.
4 Kolehmainen J, Black GC, Saarinen A *et al*: Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. *Am J Hum Genet* 2003; 72: 1359–1369.
5 Velayos-Baeza A, Vettori A, Copley RR, Dobson-Stone C, Monaco AP: Analysis of the human VPS13 gene family. *Genomics* 2004; 84: 536–549.
6 Falk MJ, Feiler HS, Neilson DE *et al*: Cohen syndrome in the Ohio Amish. *Am J Med Genet A* 2004; 128: 23–28.
7 Hennies HC, Rauch A, Seifert W *et al*: Allelic heterogeneity in the COH1 gene explains clinical variability in Cohen syndrome. *Am J Hum Genet* 2004; 75: 138–145.
8 Mochida GH, Rajab A, Eyaid W *et al*: Broader geographical spectrum of Cohen syndrome due to COH1 mutations. *J Med Genet* 2004; 41: e87.
9 Seifert W, Holder-Espinasse M, Spranger S *et al*: Mutational spectrum of COH1 and clinical heterogeneity in Cohen syndrome. *J Med Genet* 2006; 43: e22.
10 Katzaki E, Pescucci C, Uliana V *et al*: Clinical and molecular characterization of Italian patients affected by Cohen syndrome. *J Hum Genet* 2007; 52: 1011–1017.
11 Balikova I, Lehesjoki AE, de Ravel TJ *et al*: Deletions in the VPS13B (COH1) gene as a cause of Cohen syndrome. *Hum Mutat* 2009.
12 Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G: Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002; 30: e57.
13 Kolehmainen J, Wilkinson R, Lehesjoki AE *et al*: Delineation of Cohen syndrome following a large-scale genotype-phenotype screen. *Am J Hum Genet* 2004; 75: 122–127.
14 Chandler KE, Kidd A, Al-Gazali L *et al*: Diagnostic criteria, clinical characteristics, and natural history of Cohen syndrome. *J Med Genet* 2003; 40: 233–241.
15 Ariani F, Mari F, Pescucci C *et al*: Real-time quantitative PCR as a routine method for screening large rearrangements in Rett syndrome: report of one case of MECP2 deletion and one case of MECP2 duplication. *Hum Mutat* 2004; 24: 172–177.
16 Bugiani M, Gyftodimou Y, Tsimpouka P *et al*: Cohen syndrome resulting from a novel large intragenic COH1 deletion segregating in an isolated Greek island population. *Am J Med Genet A* 2008; 146A: 2221–2226.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)

3.3 Next Generation Sequencing analysis

3.3.1 Materials and Methods

Patient

We selected an Italian patient (C339) of 12 years not yet screened by DHPLC and Sanger sequencing analysis. He presents developmental delay, microcephaly, sociable behaviour, mild myopia and astigmatism, joint hyperlaxily, small hands with tapering fingers and typical facial features as downslanding palpebral fissures, synophris, prominent upper incisors and grimacing at smile.

Amplicon Library preparation

Next Generation Sequence (NGS) technology was approached in order to simultaneously analyse all *COH1* gene coding regions. The procedure starts with the preparation of an amplicon library by amplifying each amplicon separately. PCR has been performed with Fusion primers designed to generate amplicons ranging in size between 200-400 bp segments (http://454.com/downloads/my454/documentation/gs-junior/method-manuals/GSJunior_AmpliconLibraryPrep-RevJune2010.pdf) (Appendix 7.1). At the 5' end, fusion primers contained an additional sequence, Multiplex Identifier (MID), that barcodes the sample (Appendix 7.1). The MID sequence was selected from a list provided from Roche. Thermal cycling was performed on an Applied Biosystems 2720 Thermal Cycler using the following cycling profile: one cycle at 95°C for 5’ followed by 35 cycles at 95°C for 30’’, at the specific annealing temperature for 30’’(Appendix 7.1), at 72°C for 30’’, followed by a final extension step at 72°C for 5’. Small DNA fragments were removed using AMPure PCR purification system (Agencourt, Beverly, MA, USA) following the

manufacturer's protocol (http://454.com/downloads/my454/documentation/gs-junior/methodmanuals/GSJunior_AmpliconLibraryPrep-RevJune2010.pdf). Amplicons were subsequently quantified using the Quant-iT™ PicoGreen® dsDNA reagent (Invitrogen). All amplicons were then pooled at an equimolar ratio. Subsequently, the sample pool was diluted to a final concentration of 1 X 10^7 PCR fragment molecules/µl.

GS Junior sequencing

The amplicon-PCR-derived fragments were annealed to carrier beads and clonally amplified by emulsion PCR (emPCR). EmPCR was performed according to manufacturer's protocol (http://454.com/downloads/my454/documentation/gs-junior/method-manuals/GSJunior_emPCR_Lib-A_RevApril2011.pdf). The beads were isolated and compartmentalized into droplets of an aqueous PCR reaction buffer in oil emulsion. Subsequently, the emulsions were broken by isopropanol to facilitate collection of the amplified fragments bound to their specific beads. The beads carrying single stranded DNA templates were enriched, counted, and deposited into the PicoTiterPlate for sequencing (http://454.com/downloads/my454/documentation/gs-junior/method-manuals/GSJunior_Sequencing-MM-RevJune2010.pdf). The 454 technology is based on pyrosequencing, a sequencing approach based on chemiluminescent detection of pyrophosphate released during polymerase mediated deoxynucleoside triphosphate (dNTP) incorporation [27-29]. During sequencing a CCD camera-based imaging assembly was used to capture the pyrosequencing-derived light signal and to collect the readout data per flow, which was then used by a Genome Sequencer-specific base-caller to generate the sequence reads.

Data analysis

Data analysis was performed using the Roche proprietary software package for the GS Junior system. Image acquisition, image processing and signal processing were performed during the run. Post run analysis was conducted using GS Amplicon Variant Analyzer (AVA) 2.5p1 (http://454.com/downloads/my454/documentation/gs-junior/software-manual/454_Sequencing_Software_Manual_v2.5p1_PartD.pdf). The AVA application computes the alignment of reads from Amplicon libraries obtained on the GS Junior Instrument, and identifies differences between the reads and a reference sequence. In this study, amplicon nucleotide sequence reads were aligned to the Human Mar. 2006 (hg18) assembly genomic sequence of *COH1*. The AVA software identifies all nucleotide variants, and provides read counts and frequencies. Variations are also displayed graphically with a histogram indicating positions of variation.

Sanger sequencing

Direct sequencing of the purified PCR products, obtained with the same primers and PCR conditions (Appendix 1) described for amplicon library preparation, was performed in both directions (PE Big Dye Terminator Cycle Sequencing Kit) on an ABI Prism 310 genetic analyser (PE Applied Biosystems, Forest City, CA, USA) and analyzed with the Sequencer software.

3.3.2 Results

Sequencing output of patient C339 was visualized in a table containing gene name, type of variant, percentage of forward and reverse sequences containing the variant with specification of the numbers of passed filter sequences (Appendix 7.2).

The total length of *COH1* targeted regions was covered. The percentage of amplicons covered at >100-fold depth (number of reads) was 76% and the mean depth of coverage was 232.

A total of 122 variants were detected among the targeted regions of the *COH1* gene. Sanger method was used in order to distinguish real changes from false positives. Four of the variants, detected in 39-54% of Forward and Reverse sequences, were confirmed by Sanger sequencing (Tab. 5, Fig. 6). All of them are intronic substitutions, except c.10294G>A, a missense mutation changing a Glycine to an Arginine within exon 56 (Fig. 6). In all 4 cases, the variants were present in heterozygous state. Review of the literature, public databases of known sequence variants (dbSNP, http://www.ncbi.nlm.nih.gov/snp) and personal data revealed that three variations were already known as polymorphisms (rs1487024, rs3451354 and rs6468694) including the exonic one [15] (Table 5). The other intronic variant (IVS60+18C>T) was identified in this study for the first time (Table 5). In all 4 cases, the variants were present in heterozygous state.

Variants present in PolyT stretches (19) (Fig. 7), variants identified with low percentage (both strands <25%) (72) or variants with highly unbalanced percentages (one strand <5%) (26) were not confirmed by Sanger sequencing.

EXON/ INTRON	VARIANT	SNP CODE	REFERENCE	PATHOGENICITY
Intr30	IVS30-85C>T	rs1487024	-	SNP
Intr55	IVS55+24G>A	rs3451354	-	SNP
Ex56	c.10294G>A	rs6468694	Seifert et al, 2008	SNP
Intr60	IVS60+18C>T	-	-	Unknown

Table 5. Variants identified by 454 Roche technology and confirmed by Sanger sequencing. The reference sequence of *COH1* gene is according to UCSC Genome Browser, http://genome.ucsc.edu, on Human March 2006 Assembly, hg18; NM_017890.

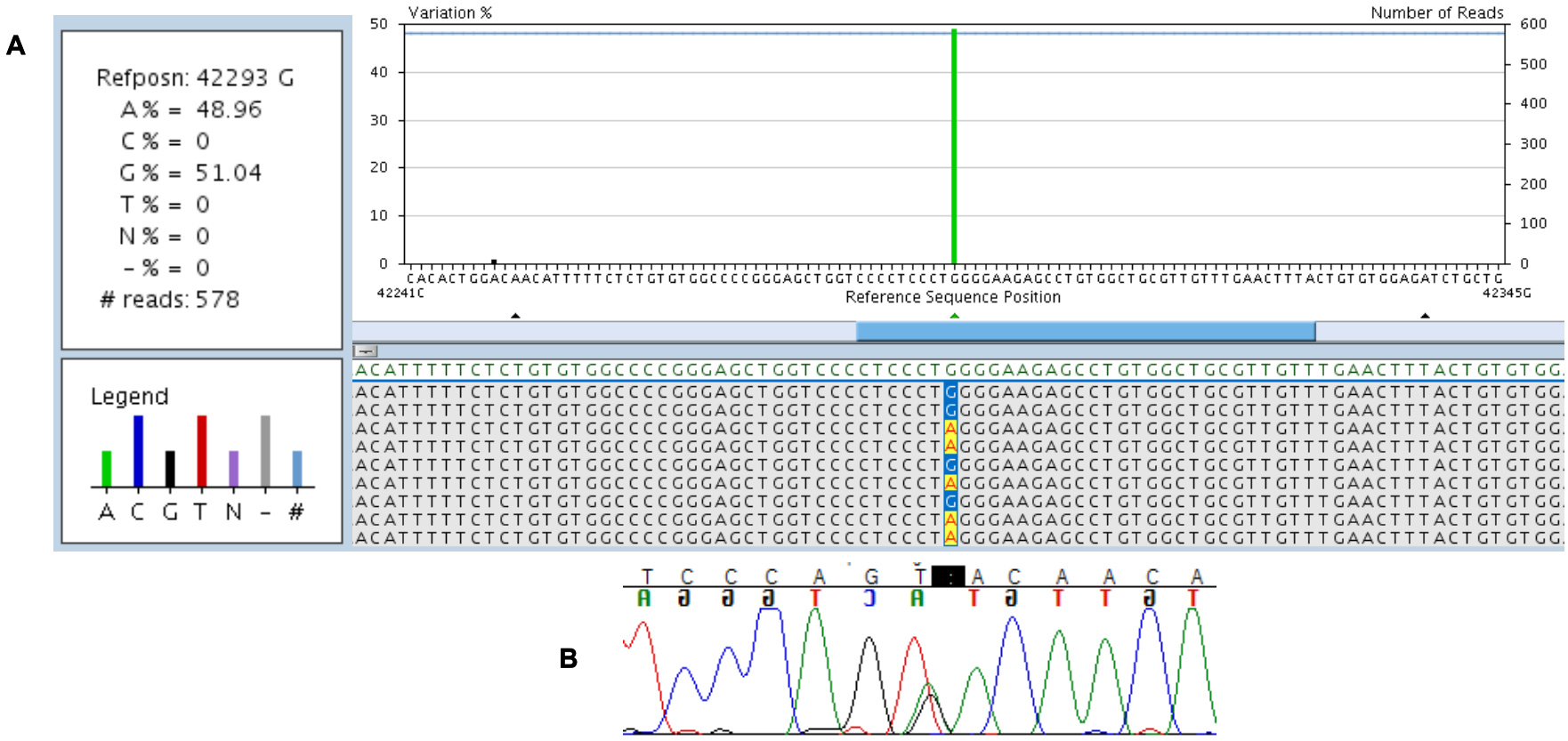

Fig. 6. Detection of the exonic variant c.10294G>A (p.G3407R) in patient C339. A) NGS results. Screenshot from the GS Amplicon Variant Analyzer Software. In the upper panel, a green histogram corresponding to an "A" (see legend on the left) represents the frequency of the mutated base found in that position; in the lower panel, alignment of the reads from both strands shows both wild-type and mutated sequences. On the left, position of the base in the reference sequence, percentages of the wild-type and mutated bases, and total number of reads are displayed. B) Sanger sequencing chromatogram of exon 56 confirming the presence of a G>A variant.

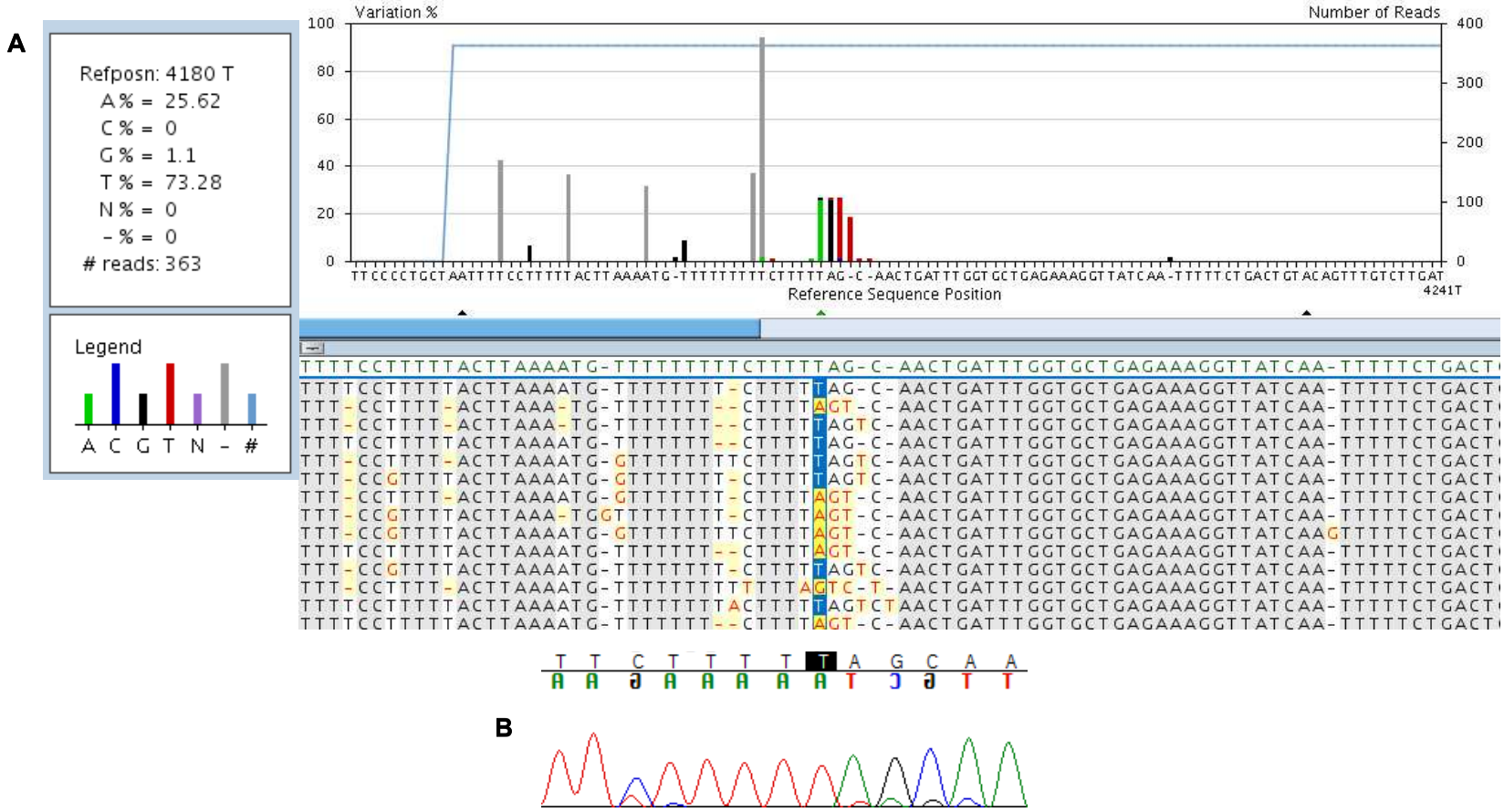

Fig. 7. Variant detected by NGS but not confirmed by Sanger sequencing. A) Screenshot from the GS Amplicon Variant Analyzer Software showing the detection of multiple variants within a polyT strech in exon 6. B) Sequence chromatogram of exon 6 showing the absence of variations.

3.3.3 Discussion

In this study, Next Generation Sequencing technology (Roche 454 GS Junior Sequencing platform) was employed for the first time for the molecular diagnosis of Cohen syndrome. Until now, *COH1* point mutation analysis has been performed by DHPLC followed by Sanger sequencing [13]. However, considering the high number of exons in the gene and the absence of mutational hotspots, this method was time consuming (about 5-6 months for a patient) and incurs relatively high costs (more than 5.000,00 euros for a patient). In addiction, some variations could be missed by DHPLC analysis since it has a sensibility of 95-98%.

Employment of the 454 GS Junior Sequencing platform allowed the simultaneous screening of all the exons of the *COH1* gene. After optimisation of the PCR conditions for each amplicon, only six annealing temperatures were established (Appendix 7.1), reducing the time of the amplification step. In total, the experiment has been completed in six working days and the cost was about 3.000,00 euros.

The NGS method allowed to disclose 122 variations within the *COH1* gene (Appendix 7.2). Four substitutions were confirmed by Sanger sequencing (Tab. 5). Three of them consist in intronic variants probably without effects on the protein. However, IVS60+18C>T has never been reported before (Tab. 5). The analysis in a control population would better clarify its pathogenicity. The only one exonic exonic variant (c.10294G>A) has been already reported in literature and in the dbSNP as benign polymorphism (rs6468694) [15] (Tab. 5).

The absence of pathogenic mutations could be consistent with the clinical phenotype of patient C339 who doesn't fulfil all the required diagnostic criteria. In particular, he shows a very mild myopia, no retinal dystrophy and no neutropenia.

Sanger sequencing has been used to evaluate real variants and to set threshold values that could help in excluding false positive results. As expected, the heterozygous variants confirmed by Sanger sequencing were detected at a percentage of about 50% in both strands (Appendix 7.2). The other variants found in a relatively high percentage of one strand but in a much lower percentage of the other strand (26) were not confirmed by Sanger sequencing (Appendix 7.2), suggesting that they represent technical artefacts. Variants detected in a percentage <25% of both strands (72) also represent false positive of NGS technology not confirmed by Sanger sequencing (Appendix 7.2). We also found 19 variants present in intronic PolyT stretches that were not confirmed by Sanger sequencing. These false positive results could be due to the difficulty in sequencing of homonucleotide regions longer than six bases by pyrosequencing. However, these repeated regions could represent mutational hotspots and an accurate detection of possible insertions/deletions is therefore needed for a diagnostic application.

In conclusion, the first successful application of 454 GS Junior Sequencing platform for Cohen syndrome molecular diagnosis has confirmed that NGS is a fast, sensitive and relatively low cost method that could be employed for the simultaneous screening of all the *COH1* coding regions.

4. Conclusions and future perspectives

4. CONCLUSIONS AND FUTURE PERSPECTIVES

Cohen syndrome is a uncommon autosomal recessive disorder first described in the Finnish population, where the phenotype is highly homogeneous due to a founder mutation present in the majority of cases [1,4,8]. It consists of non-progressive mental retardation, microcephaly, typical facial features, hypotonia and joint hyperextensibility, retinal dystrophy or myopia and intermittent neutropenia. However, in non-Finnish and especially in young Cohen patients a high allelic and phenotypic variability prevail [5-7,13].

Several diagnostic criteria have been introduced to help the clinician in establishing a diagnosis of Cohen syndrome. Before the identification of *COH1* gene, Kivitie-Kallio and Norio were the first to propose the essential features for Cohen syndrome diagnosis on the basis of the phenotype of the Finnish population [4]. These criteria were subsequently modified by Chandler et al. to be more applicable to young patients and to patients of different origin [5]. Since the identification of *COH1* as the gene responsible for Cohen syndrome, more than 100 distinct mutations were discovered in patients with different ethnic background and new clinical features were noted. Thus, Chandler criteria were revised and patients were considered as having Cohen syndrome when six of the eight following criteria are fulfilled: developmental delay, microcephaly, typical facial gestalt, truncal obesity, sociable behaviour, joint hypermobility, ocular abnormalities and neutropenia [6].

Kolemainen criteria are now commonly used during genetic counselling to establish the clinical diagnosis of Cohen syndrome and to decide for a *COH1* mutational screening. However, since mutations have been also identified in atypical patients [12,13], molecular analysis is often indicated even without all the main clinical criteria.

Considering the dimensions of the *COH1* gene and the absence of recurrent mutations in non-Finnish patients, there are very few centres in the world that perform Cohen syndrome molecular analysis. Thus, the Medical Genetics Unit of Siena has become, in the last years, an international reference centre for the molecular analysis of the *COH1* gene.

In our laboratory, point mutation analysis of all *COH1* exons was traditionally performed by DHPLC followed by automatic sequencing. Using these techniques, 96 patients with a clinical diagnosis or suspect of Cohen syndrome have been analysed. A total of 21 different point mutations were found in 22 patients from 20 families. They include 6 frameshift, 3 splicing, 11 missense mutations and one complex rearrangement (Results 3.1). All except three missense mutations have never been reported before (Results 3.1).

Among the group of mutated patients, three cases present alterations in both alleles, while in the others only one point mutation was identified. Since Cohen syndrome is a recessive condition, one homozygous mutation or two mutations in compound heterozygousity have to be found in order to confirm the clinical suspect. It is therefore possible that a rearrangement not detectable by traditional analysis could be present on the other allele.

In the last years, a consistent number of apparently benign copy number variations have been discovered across the human genome [30] and collected in the Database of Genomic Variants (http://projects.tcag.ca/variation/). *COH1* is located in a copy number variable region [30,31]. It was therefore possible that these *COH1* copy number variations could represent heterozygous mutations resulting in Cohen syndrome when present in both alleles or associated to a point mutation. In accordance, *COH1* multiple-exon deletions have been then demonstrated to be an important cause of Cohen syndrome by cDNA analysis, Real Time qPCR and targeted array [6,12,13,15,20,32].

Since incorporation of quantitative detection tools was mandatory for the molecular diagnosis of Cohen syndrome, we decided to develop a fast and low cost method allowing *COH1* intragenic rearrangement analysis. In 2002, Schouten et al. developed a technique allowing the relative quantification of up to 40 different nucleic acid sequences in a single reaction tube [33]. We therefore developed, in collaboration with MRC-Holland (www.mlpa.com), two MLPA assays for a rapid and low cost analysis of *COH1* rearrangements in all 62 exons of the gene. The application of this technique allowed us to identify 11 deleted and 4 duplicated alleles, confirming that large rearrangements account for a significant percentage of *COH1* mutations (42%) (Results 3.2).

Recently, both targeted and non-targeted oligonucleotide arrays have been approached for the identification of copy number variations in the *COH1* gene [20,34,35]. Targeted 8x15K array-CGH, comprising a total of 3148 oligonucleotides with a median resolution of 200bp and 20 oligonucleotides from the region surrounding the gene, was used for the first time by Balikova et al. in 2009 [20]. Since this technique was able to rapidly analyze at high resolution *COH1* deletions/duplications, targeted array was proposed as the method of choice for the identification of rearrangements in Cohen syndrome patients [20]. On the other hand, El Chehadeh et al. proposed the use of non-targeted array-CGH (244K), since it was more readily available and it was able to detect all the rearrangements identified by targeted array and MLPA [35]. In contrast to what stated by El Chehadeh et al, this study demonstrated that MLPA is easier, faster, and cheaper respect to array-CGH (Tab. 6). In fact, MLPA is a technically uncomplicated method which can be easily performed in each laboratory, since it only requires a thermocycler and capillary electrophoresis equipment (Tab. 6). Up to 96 samples can be handled simultaneously and results are available within 24 hours (Tab. 6). Another advantage is that only 20 ng of DNA

are required respect to array-CGH that needs 1 µg of DNA (Tab. 6). In addiction, respect to array-CGH, MLPA is not expensive: the price for each reaction is about 15€, with all the reagents included in the kit (Tab. 6). Moreover, MLPA does not require any specific software for data analysis (Tab. 6). (Results 3.2)

	MLPA	***Array-CGH***
Required equipment	• thermocycler • capillary eletrophoresis apparatus	• heat block • 2 water baths • hybridisation oven • scanner • 3 specific softwares for scanning, extraction and data analysis
Starting DNA amount	20ng	1µg
Cost per sample analysed	~15€	~600€
Working time	24 hours	3 days

Tab. 6. Comparison between MLPA and non-targeted array-CGH techniques.

MLPA also permitted to identify a deletion with the same exon coverage (6-16) in three Italian cases (R111, C42/C43 and C155) (Results 3.2). In patients R111 and C42/43, coming from Central Italy (Florence and Rome), the deletion was in heterozygous state, while patient C155, coming from Southern Italy (Barletta), presented a homozygous deletion even if the parents were apparently not consanguineous. A deletion involving the same exons was previously reported in the inbreeded population of a small Greek archipelago named Fournoi [14] (Fig. 8). Interestingly, patient C155 also shared a similar phenotype with the Greek homozygous patients (Results 3.2). Analysis with 8 microsatellite markers, performed on the three Italian families and one of the Greek patients, showed a common haplotype (Results 3.2). This suggests the possibility of a common origin of the deletion, due to an ancestral founder effect in the Mediterranean basin. In fact, Greek migration to Italy is a phenomenon that has been described starting

from the 8th century BC due to various reasons, including demographic crisis and the search for new commercial ports. This migration spread Greek people around the Mediterranean and Bleak Sea basins founding colonies in northern Lybia, eastern Spain, Southern France and especially in Southern Italy. In the XV century, Greeks also settled in Venice, Florence and Rome, establishing some of the strongest Greek communities of the time [36,37]. Thus, a possible hypothesis is that the deletion could have occurred in Greece and then it could have been spread in Fournoi archipelago as consequence of the inbreeding marriages. The continuous Greek migrations toward Italy could explain the presence of the same deletion in our country (Fig. 8). MLPA has therefore contributed to discover a possible Mediterranean founder mutation that, as well as the Finnish one, leads to a homogeneous and distinct phenotype characterized by similar facial gestalt, mild microcephaly and a severe visual impairment (Results 3.2).

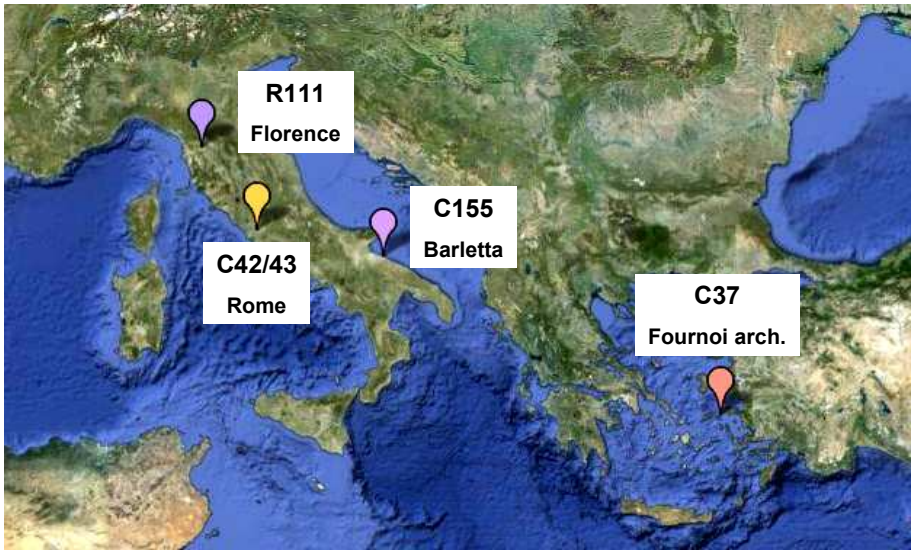

Fig. 10. Geographic distribution of the three Italian cases (R111, C42/43 and R111) and one of the Greek patients (C37) sharing the 6-16 exons deletion. The founder mutation could be originated in an ancestor from Fournoi population and it could be propagated across Mediterranean during recurrent migrations.

Moreover, for the first time, MLPA allowed to observe that even intragenic duplications can lead to the disease (Results 3.2). In fact, we found three different duplications (patients C8, C91 and C164/268) spanning exons 4-13, 20-30 and 57-60 respectively. In one case (C164/268) we performed Long Range PCR to characterize the rearrangements and to predict the effect on the protein. The characterization of the initial breakpoint showed that the duplication interrupts exon 61, causing a frameshift and a truncated protein product (Results 3.2).

Point mutation screening associated with MLPA analysis allowed to find two mutated alleles confirming the clinical diagnosis of Cohen syndrome in a total of 14 patients. All these patients show a strongly suggestive phenotype, with the fulfilling of at least 6 of the Kolehmainen criteria (Results 3.1 and 3.2). Moreover, the combined analysis lead us to identify two missense variations in one case and only one missense variation in 12 cases (Results 3.1). In these cases, it is more difficult to determine the pathogenicity of the nucleotide variation. Using Polyphen, conservation alignment programs and literature data, we could predict that four variations were probably pathogenetic (p.N204S, p.N2993S, p.H520R and p.T1068I). In these cases, the other mutated allele could have been missed because the mutation is located in an uncoding region not included in the present analysis (mutations in regulatory regions or deep intronic splicing variations) (Results 3.1). The other seven are probably polymorphisms and the analysis in a control population from the same geographic background would help to assess their pathogenicity.

Until now, DHPLC analysis followed by Sanger sequencing has been the method of choice for *COH1* point mutation analysis [13]. However, this protocol is time-consuming and incurs relatively high costs, due to the large number of exons to be investigated and to the

absence of mutational hot-spots. Usually, 5-6 months are necessary to complete Cohen syndrome molecular analysis, and the cost for a single patient is approximately 5.000 euros.

NGS technology therefore represents a very interesting tool to improve Cohen syndrome molecular diagnosis. This technique, using a Roche 454 platform, has been recently optimized in our laboratory for Alport syndrome molecular diagnosis [38]. For Cohen syndrome, we set up a specific protocol enabling the simultaneous detection of sequence variations in the 62 exons (Results 3.3). The pilot analysis by NGS in the first Cohen patient (C339) demonstrated that the molecular diagnosis can be completed in 6 working days at the cost of 3.000 euros (Results 3.3). Further reduction in time and costs will be achieved by using primers bar-coded by different MID sequences, allowing to run multiple samples in the same experiment.

In conclusion, this study has important consequences for the molecular diagnosis of Cohen syndrome, indicating MLPA and Next Generation Sequencing as the methods of choice for the analysis. In fact both techniques allowed to save time and costs respect to the other methods presently reported [20,34,35]. Improvements in accuracy and ease of data analysis in NGS will permit to further improve the diagnosis in the future.

5. References

5. REFERENCES

1 Cohen MM, Jr., Hall BD, Smith DW, Graham CB, Lampert KJ: A new syndrome with hypotonia, obesity, mental deficiency, and facial, oral, ocular, and limb anomalies. *J Pediatr* 1973; 83: 280-284.

2 Carey JC, Hall BD: Confirmation of the Cohen syndrome. *J Pediatr* 1978; 93: 239-244.

3 Norio R, Raitta C, Lindahl E: Further delineation of the Cohen syndrome; report on chorioretinal dystrophy, leukopenia and consanguinity. *Clin Genet* 1984; 25: 1-14.

4 Kivitie-Kallio S, Norio R: Cohen syndrome: essential features, natural history, and heterogeneity. *Am J Med Genet* 2001; 102: 125-135.

5 Chandler KE, Kidd A, Al-Gazali L *et al*: Diagnostic criteria, clinical characteristics, and natural history of Cohen syndrome. *J Med Genet* 2003; 40: 233-241.

6 Kolehmainen J, Wilkinson R, Lehesjoki AE *et al*: Delineation of Cohen syndrome following a large-scale genotype-phenotype screen. *Am J Hum Genet* 2004; 75: 122-127.

7 Hennies HC, Rauch A, Seifert W *et al*: Allelic heterogeneity in the COH1 gene explains clinical variability in Cohen syndrome. *Am J Hum Genet* 2004; 75: 138-145.

8 Kolehmainen J, Black GC, Saarinen A *et al*: Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. *Am J Hum Genet* 2003; 72: 1359-1369.

9 Horn D, Krebsova A, Kunze J, Reis A: Homozygosity mapping in a family with microcephaly, mental retardation, and short stature to a Cohen syndrome region on 8q21.3-8q22.1: redefining a clinical entity. *Am J Med Genet* 2000; 92: 285-292.

10 Falk MJ, Feiler HS, Neilson DE *et al*: Cohen syndrome in the Ohio Amish. *Am J Med Genet A* 2004; 128: 23-28.

11 Mochida GH, Rajab A, Eyaid W *et al*: Broader geographical spectrum of Cohen syndrome due to COH1 mutations. *J Med Genet* 2004; 41: e87.

12 Seifert W, Holder-Espinasse M, Spranger S *et al*: Mutational spectrum of COH1 and clinical heterogeneity in Cohen syndrome. *J Med Genet* 2006; 43: e22.

13 Katzaki E, Pescucci C, Uliana V *et al*: Clinical and molecular characterization of Italian patients affected by Cohen syndrome. *J Hum Genet* 2007; 52: 1011-1017.

14 Bugiani M, Gyftodimou Y, Tsimpouka P *et al*: Cohen syndrome resulting from a novel large intragenic COH1 deletion segregating in an isolated Greek island population. *Am J Med Genet A* 2008; 146A: 2221-2226.

15 Seifert W, Holder-Espinasse M, Kuhnisch J *et al*: Expanded mutational spectrum in Cohen syndrome, tissue expression, and transcript variants of COH1. *Hum Mutat* 2009; 30: E404-420.

16 Tahvanainen E, Norio R, Karila E *et al*: Cohen syndrome gene assigned to the long arm of chromosome 8 by linkage analysis. *Nat Genet* 1994; 7: 201-204.
17 Velayos-Baeza A, Vettori A, Copley RR, Dobson-Stone C, Monaco AP: Analysis of the human VPS13 gene family. *Genomics* 2004; 84: 536-549.
18 Rampoldi L, Dobson-Stone C, Rubio JP *et al*: A conserved sorting-associated protein is mutant in chorea-acanthocytosis. *Nat Genet* 2001; 28: 119-120.
19 Seifert W, Kuehnisch J, Maritzen T, Horn D, Haucke V, Hennies HC: The Cohen syndrome-associated protein, COH1, is a novel, giant Golgi matrix protein required for Golgi integrity. *J Biol Chem*.
20 Balikova I, Lehesjoki AE, de Ravel TJ *et al*: Deletions in the VPS13B (COH1) gene as a cause of Cohen syndrome. *Hum Mutat* 2009.
21 Murphy AM, Flanagan O, Dunne K, Lynch SA: High prevalence of Cohen syndrome among Irish travellers. *Clin Dysmorphol* 2007; 16: 257-259.
22 Douzgou S, Petersen MB: Clinical variability of genetic isolates of Cohen syndrome. *Clin Genet*; 79: 501-506.
23 Sunyaev S, Ramensky V, Bork P: Towards a structural basis of human non-synonymous single nucleotide polymorphisms. *Trends Genet* 2000; 16: 198-200.
24 Sunyaev SR, Lathe WC, 3rd, Ramensky VE, Bork P: SNP frequencies in human genes an excess of rare alleles and differing modes of selection. *Trends Genet* 2000; 16: 335-337.
25 Adzhubei IA, Schmidt S, Peshkin L *et al*: A method and server for predicting damaging missense mutations. *Nat Methods*; 7: 248-249.
26 Zou M, Baitei EY, Alzahrani AS *et al*: Mutation prediction by PolyPhen or functional assay, a detailed comparison of CYP27B1 missense mutations. *Endocrine*; 40: 14-20.
27 Nyren P, Pettersson B, Uhlen M: Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay. *Anal Biochem* 1993; 208: 171-175.
28 Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyren P: Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem* 1996; 242: 84-89.
29 Ronaghi M, Uhlen M, Nyren P: A sequencing method based on real-time pyrophosphate. *Science* 1998; 281: 363, 365.
30 Sebat J, Lakshmi B, Troge J *et al*: Large-scale copy number polymorphism in the human genome. *Science* 2004; 305: 525-528.
31 Redon R, Ishikawa S, Fitch KR *et al*: Global variation in copy number in the human genome. *Nature* 2006; 444: 444-454.
32 Taban M, Memoracion-Peralta DS, Wang H, Al-Gazali LI, Traboulsi EI: Cohen syndrome: report of nine cases and review of the literature, with emphasis on ophthalmic features. *J Aapos* 2007; 11: 431-437.
33 Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G: Relative quantification of 40 nucleic acid sequences by

multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002; 30: e57.

34 Rivera-Brugues N, Albrecht B, Wieczorek D *et al*: Cohen syndrome diagnosis using whole genome arrays. *J Med Genet*; 48: 136-140.

35 El Chehadeh-Djebbar S, Faivre L, Moncla A et al: The power of high-resolution non-targeted array-CGH in identifying intragenic rearrangements responsible for Cohen syndrome. *J Med Genet*.

36 Harris J: Being a Byzantine after Byzantium: Hellenic identity in Reinassance Italy. In: *Kambos: Cambridge papers in Modern Greek,* Vol 8. 2000: 25-44.

37 Harris J, Porphyriou H: The Greek Diaspora: Italian port cities and London, c.1400-1700. In: Calabi D, Christensen ST, eds. Cities and cultural transfer in Europe *Cambridge* 2007: 65-86.

38 Artuso R, Fallerini C, Dosa L *et al*: Advances in Alport syndrome diagnosis using next-generation sequencing. *Eur J Hum Genet*.

6. AKNOLEDGEMENTS

At the end of this thesis and of my PhD, there are many people that I have to thank…

First of all, I would like to thank Prof. A. Renieri for the opportunity to attend the Medical Genetics laboratory of Siena for five years, an experience that allowed me to grow up both personally and professionally. There I learn a lot of things, from correctly handling a pipette to how to write a paper, and part of the merit goes to my teachers during these years, Dr I. Meloni e Dr I. Longo.

A special thank you is for Prof. F. Ariani, the tutor of this thesis, who patiently revised and corrected this work allowing it to have a better form, and for Dr R. Artuso, my teacher during the NGS experience, who always gave me precious help and suggestions.

I would like to thank two colleagues and friends who greatly contributed to Cohen syndrome diagnosis, Dr F. Scionti e Dr Rossella Tita. I would not have done it without you! Then, I can't forget about the contribution of Dr E. Katzaki, because all started from her.

A very affectionate thank you goes to all the girls of the lab, who shared with me time, work, fun, satisfactions or desperation. Thanks for your support, assistance, or just friendship.

I would like to thank my parents and my sister, who have always been interested in what I do, giving me support and advices. And finally a very big thank you to my husband, Rossano, the one who more than all patiently stand my mood swings, my complaints and my tiredness, always waiting for me at home with an encouraging smile.

So thank you all…because this thesis also belongs to you.

7. Appendix

7. APPENDIX

7.1. Table of NGS primers.

COH1 primers designed for 454 Roche technology and corresponding annealing temperatures are reported. MID sequences are indicated in red.

Annealing (°C)	Primers	Exon
52	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTATTCAAATATGATTATACCT 3'	38
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTAAAATGCAACCACATGCAC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTTGGGAAATTTTGGTTATTGAA 3'	41
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTATGCCTGTTTTGCCAAGTC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTGTTTTTGTTGGCACGTTTGG 3'	44
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTGGGTCATGAACACAAAGAAAC 3'	
56	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTTTCTCTTTTAAGAAAGGAATTTGT 3'	3
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGTTGGCATATTTGAAAAGGTAA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTAAACATGCGTTTGTTGGTGT 3'	7
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAAAAACAATCAGCCCATCTCA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTGGGAGGAAAAATTTTAAAGGA 3'	8
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAGAGATATAAACAGCAAGAG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTAAAAATGAGAGAAGAGCGATTTTA 3'	13
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAACAAATGGCAAAAGATTCAAAA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCGCGAGTGCTTGTTTACTG 3'	35
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTTTTTGCACGGAATGTCAAG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCAGAGTGACCCAGGGAAGAG 3'	47
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCCCCAGTGCAAGGTTACTTT 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTATGAATTGCAGGGAAAATGG 3'	57
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAAGGAGTGAAGGCATTATTA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTGCCATAATGAAGTCTGTTTTTG 3'	46
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTCTCTTCCCTGGGTCACTCT 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTAAAGCATTAATGATCTCTTTTTGAGA 3'	54
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAGGCAACTGACAAAACCTCA 3'	
58	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTGAATCTTAAAATTGGCCTTGTT 3'	9
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAAAATAAGAAACAAACATCTTAC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTTGGTCATCCTTTGTGTTGTT 3'	12
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTCTGAGGAAGGCTTTGTTTCA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTGGAGAACTAATTCATTTTTCAGTTG 3'	14
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAAAAATCCATATATGAGGGAAAAA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTAACAATCTTGAAAATACGTTTGG 3'	22
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCAATGCTGGGAATTAAAAAGG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTTGCATGTAAGATGTGAAAAAG 3'	26
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGGGAGGGAAGAAAGAAAAGG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTGCCCCAGTTTATTCATCAG 3'	31
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTCACCAACTATTCTGGCAATG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTGAAATGGCTAATAATCTTTTAC 3'	32
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGCTGAAACTGTGGCAATCC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTTTCCCCTTTGTCATGTTCC 3'	39

	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAGCCCTTTCCCACAATGTTT 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCCATGGAGGGATATTTGGTT 3'	50
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAAGGCAGATTCATTTTAGACAG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTGGTGGGGACTCATTTTCAAT 3'	59
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTGTCTGCCATGTGAAAGGTT 3'	
60	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCCGCTGGAGTTTTTCTCTG 3'	2
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGATTACCGTCTAAACAAGCTG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCAGTTTCTCTTTGGGAAATATGC 3'	5
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGGCCCAAAGAAGGGTTAAGA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTTTTATTTTGTTTAACAGGCTAATGG 3'	11
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCACAAGTATTCATACAATGTGTCCA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTGTTTGCATTCCTCCACATCC 3'	20
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAGGGGTAAGGTGGGGAGAC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTGCCATGTTTAACCTTTGCTG 3'	21
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGAAATTGTGGATGCTAACTTGC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCTGGCGAAGATGTTAGGTG 3'	23
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCACCAACAAATCTAAAAATGTCC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTGTTTGTGTGCTTCCCTATTCC 3'	24
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCCTGATCTTGATGGCCAAG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCATATCCAGCATGCATTTGTC 3'	27
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTCTCAATATGTCCGAGGTATGC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCAAGCAAAACTACCAAAGACC 3'	29_2
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCCACAAAGCACAACCAACTG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTGGAGCTTGTCAGAGAAGTAAAGG 3'	36_1
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTGGAAGATCATCCTTATTCACC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTGCACACAGTTTGGCACATAG 3'	36_2
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCCTCCCCTTTTCTGTGACT 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCCAACCAAGCAAGACGACTC 3'	37
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTATGGCCAATACCCACAGAC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCCATGTTGGCTTTCGTATCC 3'	43
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCAACCTTAACAATGGCTGTTTG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTAATGAAGTCTGAATTGATGAAGC 3'	45
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAATTTCAATGGTAAAGCTGCTC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCCCAGATCATCCACACACTG 3'	48
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCATGCACACACCACTTATGG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCTTGGATGTGGTTTTTGGAAC 3'	49
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCCTTGAATATCAAAGAACATTCC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTACTGTCCTGGCAGACAATTAG 3'	51
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCAAAAGAATCGTTATGATTTGC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCATTTGACTTGATTCTCTTCC 3'	53
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTTTTCACAGGATCACAGGTAAC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCCAATTTCCACTTTGCTGTC 3'	56_2
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGGCGTGCTGTGTGACCTG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCGCTGCCTCTGAAACTACTG 3'	58
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCTCTGAGCATTTCCTGGTCA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTGGATGGCTCTGAACAGATGA 3'	60
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGAGAGCAAGCCAGGAGCTAT 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTGGGTAACTGGATTGGTGAGG 3'	61
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTTGCCCTGTTTTCTGAAACG 3'	

62	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCTGGCTAGCAGAACGCAAGA 3'	1
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCGGATCTGCCACCATAGAAC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTATAGTTGCCTTTCCCCTGCT 3'	6
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTGATTAGCATCCACAAGAGTCAA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTTCTACAAGGAAAGCCTCTAACA 3'	10
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCAGGTGAAACACAGGGAGAAA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCCACTTTTAAAATATAAAGCGAACA 3'	15
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAGCCAAACCCATGCTAACAA 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCAATCGGCTAGTACATTTGCAG 3'	18
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGCTCCTTGTAAAACGCAAGG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTAAATCAAACATTCTCAAGTGACTC 3'	19
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTGAAAAAGCACAACTGCAGAG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTGATCAGTGATGCATTGGTAGA 3'	25
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCCATGATCTGGAAATCTGTGTT 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCACCTTCTTTGCTCCTGTCC 3'	28
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTCAATTCCTTCTCTGGGTTCC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTTTAGTTCAGTGCAGTGTTGTG 3'	28B
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCAAAACAGGAGCAGTTTCAAAG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCTGAGGTCATTTTCTTCTTTCC 3'	29_1
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGGGCTGACCAGGAGATTTTC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTTCCATTCCCTCAAAGATTTC 3'	30
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTACAGGCCACACTAAGTTATTGAC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCATCTGAAAGCTATCATGTTTCT 3'	33
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCAGAATAATTAAGCAGAAAGCAGT 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCTCCCAGAAGTTGATTCAGATG 3'	34_2
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTAGCACTCCAGTAAGTTTGAAGG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCCTTGTTTGTTTTCTCAACAG 3'	42
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTACTGTAAACTGTTATGGAATGTGC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCTGTAGCTAATGTGCTCTCTGC 3'	52
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTGACAAGCAAAGAAGAACAGAC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTCCAGGAATTGTTTGTGGTATTG 3'	55
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGGATTACATGATAATTAACCAAAAGC 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTGTTCTTGTCCCTTTCTCCTC 3'	56_1
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCTCCCTGGCAGACTAAGACAG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTACAGCAGGTTGGCTGGTCAC 3'	56_3
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTCACAGACCCTCAGCTTGTG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCTCGTAAGGGTCTGGCAAC 3'	62
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTCAAGGGCAATGCAGTGAGAC 3'	
65	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTTCTCCTGTAGCTACCATAAACTGC 3'	4
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTGGTTTTATTGATCCCATTAGAAG 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTAACATTGTTTATATGACACTTGGC 3'	34_1
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTGGAGATTTAATGAACTCTTGTCTGT 3'	
	5' CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGTAAAGCATAGTTTCTATTTGCAACTTC 3'	40
	5' CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGTTGTGGTTCTCATCTGTTGTTCA 3'	

7.2. 454 GS Junior Sequencing output for patient C339.

The table indicates the type of variant identified with the position in the reference sequence, the percentages of Forward and Reverse sequences in which the variant was found and the number of reads of Forward and Reverse strands (in parentheses). Variants are ordered on the basis of the percentage of sequence variants.

Variant	% F SEQUENCE (# of SEQ)	% R SEQUENCE (# of SEQ)
40563:A/C	99(91)	0(32)
13631-13633:TTT/---	48,48(33)	80(35)
13588-13590:TTT/---	21,45(33)	80(35)
13592:T/C	15,39(33)	74,29(35)
41303:G/A	57,14(70)	55,35(91)
13636:G/C	42,42(33)	68,57(35)
17474:T/C	70(20)	45,16(31)
25801:T/C	59,32(118)	0(19)
25803:A/T	58,47(118)	0(19)
46590:C/T	46,27(218)	54,04(161)
42293:G/A	49,14(291)	49,18(287)
13594:C/T	18,18(33)	77,14(35)
25806:G/A	55,08(118)	0(119)
22800:C/T	49,05(74)	39,29(18)
13587-13590:TTTT/----	9,09(33)	80(35)
13591:T/A	94,34(33)	0(35)
13587-13591:TTTTT/-----	0(33)	80(35)
13630-13633:TTTT/----	9,09(33)	48,57(35)
20930:A/G	0(6)	34,29(35)
5670:A/T	0(28)	48,48(33)
5671:T/A	0(28)	48,48(33)
4180:T/A	3,43(204)	54,09(159)
4181:A/G	3,43(204)	54,09(159)
4182:G/T	3,43(204)	54,09(159)
13586-13591:TTTTTT/------	0(33)	48,57(35)
5699:T/G	0(28)	42,42(33)
18750:T/G	0(65)	39,13(69)
13633:T/A	33,33(33)	0(35)
20943:A/T	0(6)	17,14(35)
48182:T/G	0(17)	15,12(86)
48185:A/T	0(17)	15,12(86)
48186:T/A	0(17)	15,12(86)
2683:T/G	0(212)	27,18(183)
20944:T/A	0(6)	14,29(35)
30589:T/G	0(97)	65(20)
13357:T/C	0(33)	20(35)

13358:C/G	0(33)	20(35)
33645:T/G	0(182)	18,27(208)
28392:T/A	0(84)	43,48(23)
28393:A/G	0(84)	43,48(23)
31728:T/G	18,05(51)	0(48)
31729:G/T	18,05(51)	0(48)
13587-13592:TTTTTT/------	0(33)	17,14(35)
13595:A/G	18,18(33)	0(35)
13596:G/T	18,18(33)	0(35)
4166:T/G	0,49(204)	19,50(159)
26855:T/C	4(50)	13,33(45)
13594:C/A	15,15(33)	0(35)
40572:T/A	10,29(91)	0(32)
31388:T/G	0(51)	14,58(48)
30601:T/A	0(97)	40(20)
30603:A/C	0(97)	40(20)
30604:C/T	0(97)	40(20)
19729:T/A	26,15(65)	0(189)
40240:A/T	5,49(91)	9,38(32)
560:G/A	0(107)	13,40(97)
4151:T/G	0(204)	14,47(159)
561:A/C	0(107)	12,37(97)
562:C/G	0(107)	12,37(97)
13332:T/G	0(33)	11,43(35)
35461:T/G	0(107)	12,37(97)
4383:T/A	10,29(204)	0(159)
30585:T/C	0(97)	54,55(11)
46507:A/C	0(218)	13,04(161)
22779:A/G	0(74)	28,18(18)
22781:G/T	0(74)	28,18(18)
38295:T/G	0(146)	8,41(226)
40241:G/A	0(91)	19,15(32)
40242:T/G	0(91)	19,15(32)
13586-13592:DEL(7)	0(33)	8,57(35)
13629-13633:TTTTT/-----	9,09(33)	0(35)
23129:T/G	5,41(74)	0(18)
5221:G/A	7,07(165)	0(135)
5224:A/G	7,07(165)	0(135)
16428:T/G	0(123)	7,07(120)
3468:T/C	0(124)	17,24(29)
3469:C/G	0(124)	17,24(29)
16691:C/A	6,09(123)	0(120)
16693:T/C	6,09(123)	0(120)
4968:T/C	0(165)	5,19(135)
34699:A/G	4,22(157)	0(121)

21472:G/A	6,56(61)	0(126)
15118:G/A	3,18(283)	0(194)
15119:T/G	3,18(283)	0(194)
15120:A/T	3,18(283)	0(194)
21268:A/C	6,37(67)	0(161)
4969:C/G	0(165)	4,10(135)
4174:T/A	0,49(204)	3,14(159)
24565:T/C	3,03(152)	1,03(158)
33647:T/G	0(182)	3,28(208)
34417:T/C	0(157)	3,31(121)
34418:C/T	1,04(157)	2,48(121)
34698:G/C	2,55(157)	0(121)
27812:T/G	3,28(122)	0(170)
27813:A/T	3,28(122)	0(170)
5236:A/C	2,42(165)	0(135)
12674:T/G	0(199)	2,14(234)
4175:C/T	0(204)	2,52(159)
4179:T/A	0(204)	2,52(159)
4180:T/G	0(204)	2,52(159)
4181:A/T	0(204)	2,52(159)
4182:G/C	0(204)	2,52(159)
4183:C/T	0(204)	2,52(159)
29352:T/C	0,41(243)	2,13(173)
42249:A/G	1,37(291)	0,35(287)
15846:A/G	1,19(254)	1,24(239)
15942:T/G	1,19(254)	1,24(239)
42398:G/C	0,34(291)	1,05(287)
20343:T/C	1,16(395)	1(501)
20316:T/C	1,27(395)	0(501)
37809:A/C	42,34(31)	0(0)
37810:T/A	32,26(31)	0(0)
37811:C/T	32,26(31)	0(0)
37817:A/T	13,30(31)	0(0)
37819:C/A	13,30(31)	0(0)
39044-39051:DEL(8)	0(0)	73,33(15)
39045-39049:TTTTT/-----	0(0)	93,33(15)
39050:T/A	0(0)	27,07(15)
39051:T/G	0(0)	27,07(15)
39052:A/T	0(0)	27,07(15)
44792:C/A	0(0)	7,09(127)
44793:A/C	0(0)	6,30(127)

7.3. Curriculum vitae and list of publications

Name: Veronica Surname: Parri

Date of Birth: 19/11/1982 Place of Birth: Siena

Studies:

04/2007 – Degree in Molecular Biology at the University of Siena

09/2004 – Bachelor Degree in Biological Sciences at the University of Siena

Experience:

During my PhD experience at the Medical Genetics Unit of Siena, I performed molecular screening on the DNA of Cohen syndrome patients, I learned cytogenetics techniques employed in prenatal diagnosis and how to do a kariotype analysis, I participated in the characterization of ACSL4, a protein involved in mental retardation, and I dealt with lymphoblastoid cell lines.

Publications:

High frequency of *COH1* intragenic deletions and duplications detected by MLPA in patients with Cohen syndrome.

Parri V, Katzaki E, Uliana V, Scionti F, Tita R, Artuso R, Longo I, Boschloo R, Vijzelaar R, Selicorni A, Brancati F, Dallapiccola B, Zelante L, Hamel CP, Sarda P, Lalani SR, Grasso R, Buoni S, Hayek J, Servais L, de Vries BB, Georgoudi N, Nakou S, Petersen MB, Mari F, Renieri A, Ariani F.

Eur J Hum Genet. 2010 Oct;18(10):1133-40. Epub 2010 May 12.

The XLMR gene ACSL4 plays a role in dendritic spine architecture.

Meloni I, Parri V, De Filippis R, Ariani F, Artuso R, Bruttini M, Katzaki E, Longo I, Mari F, Bellan C, Dotti CG, Renieri A.

Neuroscience. 2009 Mar 17;159(2):657-69. Epub 2008 Dec 24.

RSK2 enzymatic assay as a second level diagnostic tool in Coffin-Lowry syndrome.

Micheli V, Sestini S, Parri V, Fichera M, Romano C, Ariani F, Longo I, Mari F, Bruttini M, Renieri A, Meloni I.

Clin Chim Acta. 2007 Sep;384(1-2):35-40. Epub 2007 May 26.