MOLECULAR CHARACTERIZATION AND SPECTRUM OF MUTATIONS ASSOCIATED WITH AUTOSOMAL DOMINANT HYPERCHOLESTEROLEMIA

Maria Nicoletta D'Agostino

Dottorato in Scienze Biotecnologiche – XXVI ciclo
Indirizzo Biotecnologie Mediche
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche –XXVI ciclo
Indirizzo Biotecnologie Mediche
Università di Napoli Federico II

MOLECULAR CHARACTERIZATION AND SPECTRUM OF MUTATIONS ASSOCIATED WITH AUTOSOMAL DOMINANT HYPERCHOLESTEROLEMIA

Maria Nicoletta D'Agostino

Dottoranda: Maria Nicoletta D'Agostino

Relatore: Prof. Giuliana Fortunato

Coordinatore: Prof. Giovanni Sannia

A Silvio

INDICE

Introduction- Familial Hypercholesterolemia (FH) is a common genetic disease that is inherited in an autosomal dominant manner. FH can be caused by mutations in the *low-density lipoprotein receptor gene* (*LDLR*) in 80% of cases, in its main ligand *apolipoprotein B gene* (*ApoB-100*) in 5% of cases, in the *propotein convertase subtilin/kexin type 9 gene* (*PCSK9*), an enzyme involved in LDLR degradation, in 1-2% of cases, while the genetic alteration is still unidentified in 15% of cases.

FH leads to premature coronary heart disease(CHD), namely myocardial infarction and angina pectoris, due to elevated plasma low-density lipoprotein cholesterol (LDL-C) levels and high levels of total cholesterol (TC).

About 1 in 500 individuals in the general population is affected, making it the most common monogenic form of hypercholesterolemia. Patients with homozygous alleles for defective LDLR are very rare, the frequency is about $1:10^6$.

Three scientific groups have developed diagnostic tools for FH: The US MedPed Program (Make Early Diagnosis, Prevent Early Death), the Simon Broome Register Group in the United Kingdom , and the Dutch Lipid Clinic Network (DLCNC). The major difference between these systems is the use of different cut-offs for premature CHD.

The US MedPed Program based on validated age and sex adjusted for serum cholesterol cut-off points. The Simon Broome criteria take into account that TC and LDL-C levels differ for adults and children. The criteria also take into account of evidence of dominant transmission and the age of onset of coronary disease in the kindred. Using this approach, cases are categorized as ‘definite’ and ‘possible’. The DLCNC score is based on a family history of hypercholesterolemia, premature CHD, clinical features, LDL-C levels and DNA analysis. It uses a numeric score for each criterion, and classifies individuals as “definite”, “probable” or “possible” FH. A diagnosis is considered definite if the score is greater than 8, considered probable if the score is between 6 and 8 points, and considered possible when the score is between 3 and 5 points.

Aims- To determine the frequency and spectrum of mutations causing FH in patients attending the different Italian lipid clinics. To investigate genotype-phenotype correlations in FH carrying different mutations of the LDL-R gene. To identify the presence of a likely polygenic cause due to the inheritance of LDL-C-raising SNPs. In addition, to evaluate which of diagnostic criteria DCLN or Simon Broome is more accurate to detect patients with mutations in our population.

Results- Mutations were found in 214/322 subjects with a mutation rate of 66%. Out of 214 mutated subjects about 95.3% were carriers of *LDLR* mutations, 2.8% of *APOB* mutations and 1.8% of *PCSK9* mutations. In the LDLR gene we found 58 different mutations of which 5 were novels

(c.102C>A; c.892A>G; c.1277T>G; c.694+1G>C c.1070_1071dupAG). Regarding the APOB gene we identified four different mutations of which two novels (p.Val3306Ile and p.Trp3633Arg).
In the PCSK9 gene we found four different variations of which two novels (p.Pro331Ala, p.Arg499Hys). The gradually increase of the TC and LDL-C among patients with different types of mutation shows that the type of the LDLR mutation influences the lipid profile. Patients with radical mutations show a worse lipid profile than missense carriers allowing a prognostic evaluation for physicians. In a proportion of patients with the FH phenotype but without mutations in the main candidate genes, there is a likely polygenic cause due to the inheritance of LDL-C-raising SNPs which increases LDL-C concentration in patients.
In addition, we examined the reliability of the Dutch Lipid Clinic Network (DCLN) score and Simon Broome criteria, to identify patients with a high or low probability of carrying an FH-causing mutation. The DCLN criteria was more accurate than Simon Broom showing very high sensitivity and specificity to detect patients with mutations.
Conclusions- The results of this study represent an update of FH genetic background in an Italian population of patients from southern regions.
These data enlarge the spectrum of mutations causing FH. The correlation between mutation types and lipid profile underlines importance of genetic screening as a prognostic tool. The comparison of the Dutch Lipid Clinic Network (DCLN) score and Simon Broome revealed that genetic screening is also useful to confirm the diagnosis, especially in patients with an uncertain phenotype.

Riassunto

L'Ipercolesterolemia Familiare (FH) è una malattia ereditaria, caratterizzata da un alterato metabolismo lipoproteico, che coinvolge quasi esclusivamente la frazione delle lipoproteine a bassa densità (Low Density Lipoprotein – LDL).
L'Ipercolesterolemia Familiare (FH) è una patologia ereditaria, caratterizzata da un alterato metabolismo lipoproteico, che coinvolge quasi esclusivamente la frazione delle lipoproteine a bassa densità (Low Density Lipoprotein – LDL).
La forma più comune di trasmissione è quella autosomica dominante che nella forma eterozigote ha una prevalenza di circa 1:500 individui, e nella forma omozigote ha una prevalenza di circa $1x10^6$ individui. La malattia nella forma omozigote è caratterizzata da un fenotipo clinico con manifestazioni cliniche più severe che compaiono già in età pediatrica.
Le mutazioni nel gene codificante per il recettore delle LDL (LDL-R) sono la causa più frequente, infatti, sono responsabili di circa 80% dei casi dell'FH. Un secondo gene coinvolto è l'APOB che codifica per l'apolipoproteina B, mutazioni in questo gene, sono identificate in circa 5-10% dei casi, infine un terzo gene coinvolto è PCSK9 che codifica per la Proproteina Convertasi Subtilisina/Kexina tipo9 responsabile di FH di tipo III in circa 1-2% dei casi.
Lo studio dei tre geni responsabili della malattia è stato effettuato su 432 pazienti, di cui 322 appartenenti a singole famiglie. Sono state identificate mutazioni in 214/322 soggetti (tasso di mutazione 66.5%). Dei 214 soggetti mutati, 204 (95.3%) presentavano mutazione nel gene LDLR, 6 (2.8%) mutazione in *APOB* and 4 (1.8%) mutazioni in *PCSK9*. Mediante le tecniche di sequenziamento diretto e Multiplex ligation-dependent probe amplification (MLPA) nel gene LDLR sono state identificate 58 differenti mutazioni, di cui circa il 50% rappresentano le mutazioni più comuni nelle regioni del Sud Italia questo dato è in accordo con quanto riportato in letteratura, la distribuzione delle mutazioni nel gene del recettore delle LDL tende a formare dei cluster tipici delle varie aree geografiche. Le mutazioni presenti solo nella nostra popolazione di soggetti FH sono 15 di cui 10 già descritte in due precedenti lavori e 5 riportate per la prima volta in questo studio di tesi. Sono state inoltre identificate 2 nuove mutazioni sia nel gene ApoB che nel gene PCSK9. Le nuove mutazioni sono state valutate mediante criteri indiretti quali: l'assenza delle mutazioni in 150 cromosomi provenienti da individui normocolesterolemici dello stesso gruppo etnico; la segregazione delle varianti con la patologia negli individui della stessa famiglia, la conservazione dei residui aminoacidici mutati nelle proteine omologhe di diverse specie. Le nuove mutazioni identificate nei geni LDLR,

APOB, PCSK9 sono state ricercate anche nel database che si avvale del sequenziamento degli esomi di 6500 individui, Exome Sequencing Project (ESP). Nessuna mutazione è stata ritrovata in EPS, inoltre, mediante l'utilizzo di software bioinformatici è stato possibile predire la causatività delle mutazioni mediante analisi in silico. Le analisi statistiche sono state effettuate con l'utilizzo di PASW versione 18.0 software (SPSS Inc.). Sono state osservate differenze statisticamente significative ($p<0.0001$) tra i gruppi di pazienti con e senza mutazione sia per i livelli di colesterolo-totale (TC) sia per i livelli di LDL-colesterolo (LDL-C). I pazienti con mutazione sono stati suddivisi in 4 gruppi in base al tipo di mutazione presente missenso, radicale, assenza di mutazione o presenza di due mutazioni (omozigoti o eterozigoti composti). L'analisi statistica dei dati biochimici correlati al tipo di mutazione, ha evidenziato come i livelli lipidici aumentavano in modo significativo dai pazienti senza mutazione ai pazienti con una mutazione missenso a pazienti con una mutazione radicale fino a quelli con due mutazioni. Tale significatività è stata osservata sia nel confronto fra i singoli gruppi mediante l'uso della correzione del Bonferroni sia nel confronto di tutti i gruppi mediante il Test di Kruskal-Wallis.

I criteri diagnostici per porre diagnosi clinica di FH sono stati sviluppati, a livello internazionale, dal *MedPed Program* negli Stati Uniti (Make Early Diagnosys, Prevent Early Dead), dal *Simon Broome Register Group* nel Regno Unito e dal *Dutch Lipid Clinic Network* in Olanda.

Il criterio del MedPed Program è basato sulla valutazione della colesterolemia in rapporto all'età del soggetto e al grado di parentela con altri familiari nei quali è stata già diagnosticata la FH. In mancanza di dati dei familiari si adotta un livello di cut-off del colesterolo totale che tiene conto della distribuzione della colesterolemia nella popolazione generale. Il criterio stabilito dal *Simon Broome Register Group* tiene conto oltre che dei livelli di CT e LDL-C in relazione all'età del paziente, anche di altre manifestazioni cliniche della malattia (xantomi tendinei), dell'anamnesi familiare e della diagnosi molecolare. Il criterio diagnostico del *Dutch Lipid Clinic Network* (DCLN) attribuisce invece un punteggio compreso tra 1 e 8 a ciascun parametro di rilievo clinico per la patologia e quindi considera la diagnosi "certa" se il punteggio totale è superiore ad 8, "probabile" se il punteggio totale è compreso tra 6 ed 8 e "possibile" se il punteggio totale è tra 3 e 5. Secondo questi criteri, l'identificazione di mutazioni permette l'assegnazione di 8 punti e consente di porre diagnosi certa di FH.

Le linee guida del Consensus Statement of the European Atherosclerosis Society (EAS) sottolineano l'importanza dell'utilizzo dei criteri diagnostici riportati nei DCLN e nei Simon Broome perché dotati di una maggiore

accuratezza diagnostica. Pertanto i 322 pazienti inclusi nello studio sono stati classificati in gruppi secondo quanto riportato nei criteri diagnostici DCLN e nei Simon Broome. Sono state effettuate delle analisi statistiche di correlazione tra presenza/assenza di mutazione sia nei gruppi DCLN che nei gruppi Simon Broome. La frequenza di mutazione osservata aumentava con l'aumentare della gravità del fenotipo clinico con un più alto numero di pazienti mutati presenti nel gruppo di FH-definita secondo entrambi i criteri diagnostici. Inoltre è stata valutata la correlazione tra il profilo lipidico dei pazienti con e senza mutazione e la classificazione secondo i due criteri diagnostici. Anche in questo caso i livelli di TC e LDL-C aumentano gradualmente da pazienti senza mutazione a quelli classificati con una probabile/possibile-FH a quelli classificati con una definita-FH che presentano dei valori di LDL-C quasi raddoppiati rispetto ai pazienti senza mutazione. Per verificare quale dei due criteri diagnostici fosse più accurato nel classificare i pazienti oggetto di questo studio il DCLN score e il Simon Broome sono stati correlati.
Mediante l'utilizzo del test del chi-quadrato si è evidenziato che i due criteri nel 68% dei casi sono sovrapponibili, mostrando una concordanza del 20.8% nella classificazione di pazienti senza mutazione nel gruppo con FH improbabile, del 42.1% nei pazienti con una possibile-FH e un 5.1% di pazienti con definita-FH. Le discrepanze sono dovute alla presenza di un 16.1% di pazienti classificati come probabile/possibile-FH secondo i DCLN, come no-FH secondo i Simon Broome; da sottolineare inoltre la presenza di due pazienti classificati come definita-FH per i DCLN e come no-FH secondo i Simon Broome. Entrambi i pazienti presentano mutazione nel gene *LDLR;* la differenza nella classificazione è dovuta ad una diversa valutazione delle manifestazioni cliniche e storia personale tra i due criteri: un paziente presenta l'arco corneale e l'altro paziente presenta una placca carotidea manifestazioni cliniche non valutate nei Simon Broome ma considerati nei DCLN. Alla presenza di arco corneale viene attribuito un punteggio di 4 e alla presenza di placca carotidea viene attribuito un punteggio di 2, questi punteggi associati ai punteggi relativi ai livelli di LDL-C permettono di assegnare uno score maggiore di 8 consentendo di porre diagnosi certa di FH per entrambi i pazienti.
Inoltre, in collaborazione con il gruppo di ricerca del Prof. Humphries dell'University College London è stata valutata la presenza di una componente poligenica in un sottogruppo di 199 soggetti con e senza mutazione nei tre geni candidati. Studi recenti hanno evidenziato come dalla valutazione della clinica verso la presenza di mutazione si osservano casi in cui alcuni soggetti con mutazione non hanno una diagnosi clinica di FH oppure soggetti senza mutazione sono classificati come FH. In questo

gruppo l'assenza di mutazione può essere dovuta alla presenza di varianti in altri geni coinvolti nel pathway del metabolismo delle LDL ma non ancora identificati come associati alla malattia o alla presenza di una componente poligenica. In un recente studio Talmud et al hanno osservato come la presenza di 12 SNP riportati dalla *Global Lipid Genetic Consortium* (GLGC) in pazienti senza mutazione fossero direttamente associati ad alti livelli di LDL-C. I nostri risultati confermano quanto precedentemente riportato e mostrano che anche nella nostra popolazione il fenotipo di FH in pazienti senza mutazioni nei tre principali geni candidati potrebbe essere dovuto ad una componente poligenica dovuta alla presenza di SNP responsabili dell'aumento dei livelli di LDL-C. Questo contributo poligenico è stato osservato anche nei pazienti con mutazione.

L'analisi è stata estesa ad un sottogruppo di pazienti pediatrici (n=99) di età tra compresa 10 e 4 anni. In tale popolazione è stata osservata una differenza statisticamente significative tra i vari tipi di mutazione (missenso, radicale, assenza di mutazione o presenza di mutazione) e il profilo lipidico. E' stato osservato che i livelli di LDL-C aumentano considerevolmente tra il gruppo di pazienti aventi due mutazioni (13.1±0.8 mmol/l) rispetto al gruppo senza mutazioni (4.0±1.6mmol/l).

In conclusione, i risultati di questo studio rappresentano un aggiornamento dello studio genetico della Ipercolesterolemia Familiare nella popolazione italiana costituita da pazienti provenienti prevalentemente da regioni del Sud Italia.

Questi dati evidenziano un aumento del numero di mutazioni causative di FH e sottolineano l'importanza di un completo screening dei geni coinvolti nella patologia per identificare nuove mutazioni e per identificare soggetti portatori di due mutazioni. La correlazione tra tipo di mutazione e profilo lipidico sottolinea l'importanza dell'analisi genetica come uno strumento prognostico. La presenza di una considerevole percentuale di soggetti mutati nel gruppo no-FH sulla base dei Dutch Lipid Clinic Network (DCLN) score e Simon Broome criteria suggeriscono che l'analisi genetica è anche utile per confermare la diagnosi specialmente nei pazienti con un fenotipo incerto.

1. Introduction

1.1 *Familial Hypercholesterolemia*

Monogenic hypercholesterolemia is a lipid metabolism disorder characterized by an increase in cholesterol. Cholesterol plays an essential role in the normal function of nearly all cells and is required by all steroid-producing cells. Cholesterol, like most steroids, is a lipophilic molecule that circulates in blood bound to binding proteins that increase steroid solubility. Cholesterol is package in lipoprotein complexes referred to as high-density lipoproteins (HDL), low-density lipoproteins (LDL), and very-low-density lipoproteins (VLDL).
Monogenic hypercholesterolemia includes two genetic subtypes designated as Autosomal Dominant Hypercholesterolemia (ADH) and Autosomal Recessive Hypercholesterolemia (ARH) [1].
Familial Hypercholesterolemia (FH, OMIM # 143890) is a common genetic cause that is inherited in an autosomal dominant manner. FH can be caused by mutations in the *low-density lipoprotein receptor gene* (*LDLR*) or in its main ligand *apolipoprotein B gene* (*ApoB-100*) [2], or in the *propotein convertase subtilin/kexin type 9 gene* (*PCSK9*) an enzyme involved in LDLR degradation [3] or in the other still unidentified genes [4].
Recently, several loci have been reported with linkage studies, however to date this has not led to the identification of a specific causal gene [5].
A recessive form of hypercholesterolemia is caused by very rare mutations in the low density lipoprotein receptor adaptor protein 1(LDLRAP1) gene [6].
FH leads to premature coronary heart disease(CHD), namely myocardial infarction and angina pectoris, due to elevated plasma low-density lipoprotein cholesterol (LDL-C) levels and high levels of total cholesterol (TC) [7].
The clinical hallmark of FH is cholesterol ester accumulation in tissues manifesting tendon xanthomata and xanthelasmas involving the skin and extremities as well as the ocular orbit and the cornea (*arcus corneae*). However, a clinical diagnosis of FH is often made in the absence of tendom xanthomata, which are not usually present before the fourth decade of life. Despite its strong genetic background, FH shows a great variability in phenotypic expression in terms of the lipid profile, frequency of xanthomas, and onset and severity of CHD.
It is well documented in the literature that FH is associated with substantial excess mortality from coronary heart disease in young adults but may not be associated with a substantial excess mortality in older patients.

Men and women with FH between 20 and 39 years old have a 100-fold increased risk to develop premature CHD compared to the normal population. It is thus important to identify, as early as possible, individuals with FH to improve their prognosis by administration of appropriate therapeutic interventions and it is crucial to perform cascade screening to detect familiars with the illness, especially children, before the manifestations of clinical symptoms [8].

About 1 in 500 individuals in a general population is affected, making it the most common monogenic form of hypercholesterolemia. Patients with homozygous alleles for defective LDLR are very rare, the frequency is about $1:10^6$. The clinical phenotype is more severe for homozygotes than heterozygotes; these patients usually die of severe coronary artery plaque formation. FH patients are increased lifetime risk for cardiovascular disease (CVD) during adolescence and, if left untreated, clinical symptoms of CVD typically manifest in men in their fourth decade and in women in their fifth decade of life in heterozygous FH patients. Usually, heterozygous patients have LDL-C concentrations less than 5mmol/l, however, the range of LDL-C concentrations (5-11mmol/l) in heterozygous patients overlaps with concentrations observed in polygenic hypercholesterolemia [9]. Homozygous FH is characterized by extremely high concentrations of TC (usually in the range 15-24mmol/l) and LDL-C about 15.89mmol/l [9]. Patients with homozygous FH may have two identical mutations (true homozygote) or compound heterozygote having different mutations on each LDLR allele. Mutations in both *LDLR* alleles that result in reduced uptake and clearance of LDL-C is the most common cause of homozygous. Homozygotes are classified into two major groups based on the amount of LDLR activity: patients with less than 2% of normal LDLR activity are classified as receptor negative and patients with 2 to 25% of normal LDLR activity as receptor-defective [10].

Homozygous patients can experience serious cardiovascular events as early as childhood [11]. Children with FH, particularly the homozygous form, have a much higher risk for premature CHD. Evidence that hypercholesterolemia is associated with the atherosclerotic process from childhood and is accompanied by increased prevalence of cardiovascular risk factors, justifies the screening for dyslipidemia in children. Children are subject to fewer environmental influences in terms of their lipoprotein metabolism, and an unequivocal diagnosis of ADH is easier to establish [12] . As a consequence, identification of FH during childhood, at the time of the silent and reversible phase of the disease, seems to be of paramount importance.

1.2 *Diagnostic criteria for FH*

Clinical criteria used to identify patients with FH include: high plasma levels of total and LDL cholesterol, family history of hypercholesterolemia especially in children, physical examination, and personal and family history.The clinical phenotype of FH patients is variable. Three groups have developed diagnostic tools for FH: The US MedPed Program (Make Early Diagnosis, Prevent Early Death), the Simon Broome Register Group in the United Kingdom, and the Dutch Lipid Clinic Network (DLCNC). The major difference between these systems is the use of different cut-offs for premature CHD.
The first, is based on validated age and sex adjusted serum cholesterol cut-off points (Table 1). Different cut-off points are provided for each category for four age groups. The sensitivity and specificity of using cholesterol levels for diagnosis varies with the population involved. The authors estimated the cut-off levels of total cholesterol required to achieve 98% specificity, and showed that those levels would provide 88% sensitivity for screening first degree relatives and 54% sensitivity for screening the general population [13] [14].

Table 1. US MedPed Program diagnostic criteria for familial hypercholesterolemia

	Total cholesterol cutpoints (mmol/liter)			
	First-degree relative with FH†	Second-degree relative with FH	Third-degree relative with FH	General population
Age (years)				
<20	5.7	5.9	6.2	7.0
20–29	6.2	6.5	6.7	7.5
30–39	7.0	7.2	7.5	8.8
≥40	7.5	7.8	8.0	9.3
Diagnosis (FH is diagnosed if total cholesterol levels exceed the cutpoint)				

The DLCNC score is based on a family history of hypercholesterolemia, premature CHD, clinical features, LDL-C levels and DNA analysis. It uses a numeric score for each criterion, and classifies individuals as “definite”, “probable” or “possible” FH. A diagnosis is considered definite if the score is

greater than 8, considered probable if the score is between 6 and 8 points, and considered possible when the score is between 3 and 5 points.
If the score is below 3 points not diagnosis is made (Table 2) [15, 16]. It is also recommended that secondary causes of hypercholesterolemia (hypothyroidism, nephrotic syndrome, diabetes and medications) be excluded. The DCLNC criteria use a scoring system that is effectively similar to the Simon Broome Register criteria.

Table 2. Dutch Lipid Clinic Network diagnostic criteria for familial hypercholesterolemia

	Description
Criteria	
a	Total cholesterol concentration above 7.5 mmol/liter in adults or a total cholesterol concentration above 6.7 mmol/liter in children aged less than 16 years, or
	Low density lipoprotein cholesterol concentration above 4.9 mmol/liter in adults or above 4.0 mmol/liter in children
b	Tendinous xanthomata in the patient or a first-degree relative
c	DNA-based evidence of mutation in the *LDLR* or *APOB* gene
d	Family history of myocardial infarction before age 50 years in a second-degree relative or before age 60 years in a first-degree relative
e	Family history of raised total cholesterol concentration above 7.5 mmol/liter in a first- or second-degree relative
Diagnosis	
A "definite" FH† diagnosis requires either criteria *a* and *b* or criterion *c*	
A "probable" FH diagnosis requires either criteria *a* and *d* or criteria *a* and *e*	

The Simon Broome Register criteria, take into account that TC and LDL-C levels differ for adults and children (Table 3). The criteria also take account of evidence of dominant transmission and the age of onset of coronary disease in the kindred. Using this approach, cases are categorized as 'definite' and 'possible'. If a mutation has not been identified, the presence of tendon xanthomas is required for a definite diagnosis of FH in Simon Broome Register criteria, whereas in the Dutch scoring system the presence of xanthomas only increases the score [13, 17].

Table 3. Simon Broome Familial Hypercholesterolemia Register diagnostic criteria for familial hypercholesterolemia

	Points
Criteria	
Family history	
First-degree relative with known premature (men: <55 years; women: <60 years) coronary and vascular disease, or	
First-degree relative with known LDLC† above the 95th percentile	1
First-degree relative with tendinous xanthomata and/or arcus cornealis, or	
Children aged less than 18 years with LDLC above the 95th percentile	2
Clinical history	
Patient with premature (men: <55 years; women: <60 years) coronary artery disease	2
Patient with premature (men: <55 years; women: <60 years) cerebral or peripheral vascular disease	1
Physical examination	
Tendinous xanthomata	6
Arcus cornealis prior to age 45 years	4
Cholesterol levels (mmol/liter)	
LDLC, ≥8.5	8
LDLC, 6.5–8.4	5
LDLC, 5.0–6.4	3
LDLC, 4.0–4.9	1
DNA analysis	
Functional mutation in the *LDLR* gene	8
Diagnosis (diagnosis is based on the total number of points obtained)	
A "definite" FH† diagnosis requires more than 8 points	
A "probable" FH diagnosis requires 6–8 points	
A "possible" FH diagnosis requires 3–5 points	

However, in some cases tendon xanthomata with normal plasma cholesterol may be seen in sitosterolemia and cerebrotendinous xanthomatosis but these should be excluded with plasma phytosterol and DNA testing [18].

Furthermore**,** the Simon Broome Register criteria consider a molecular diagnosis as evidence for definite FH, the DLCNC requires that at least one

other criterion be met in addition to molecular diagnosis. These are the main differences between the DLCNC scoring system and the Simon Broome Register criteria [14].

In children, the diagnosis of FH cannot be made using the DLCNC score because children do not develop the physical signs of FH. In children and adolescents, the diagnosis of FH will be made using age and sex adjusted LDL-C levels and, where a family mutation is known, by DNA testing [19].The current approach to screening children and adolescents with a fasting lipid profile remains a targeted approach based on a family history of early coronary heart disease or elevated cholesterol levels. The absence of a positive genetic test in the parents does not exclude FH in a child with high cholesterol levels [20]. Testing of children families known to have a clear diagnosis of FH should be screened before the age of 10 years [21].

FH may be suspected in children if they have an untreated fasting LDL-C level above 3.5mmol/l; however, secondary causes of hypercholesterolemia (nephrotic syndrome and hypothyroidism) should first be excluded [22].

In children with heterozygous clinical signs such as xanthomas and corneal arcus are not pathognomonic because they appear later in life; however, if present, they are suggestive of homozygous FH. The National Institute for Health and Clinical Excellence (NICE) guidelines state that all patients with a clinical diagnosis of FH should be offered a DNA diagnostic test and referral for family cascade testing in order to identify relatives affected [17]. Since FH is an autosomal dominant disease, the probability is 50% for a child of parent with FH to inherit the disease. Indeed, the most effective strategy for diagnosing patients with FH is to screen close relatives of patients already diagnosed with FH. In families where the mutation has been identified, genetic testing should also be included in cascade screening [23].

Cascade screening should start with first–degree relatives and then be extended to second and third degree relatives. A systematic strategy for detecting index cases of FH is essential. The index case is the trigger for cascade screening, whereby new cases can be efficiently discovered. Potential index cases should be sought amongst patients aged less than 60years with CVD presenting to coronary, stroke and vascular care units [24].

1.3 *Molecular basis of FH*

1.3.1 LDL receptor gene

The human gene encoding the LDLR is situated on chromosome 19p13 with 18 exons that span approximately 45kb. The human LDLR messenger RNA (mRNA) is 5.3kb in length and encodes a protein of 860 amino acids [25].
The LDLR is a multifunctional cell-surface glycoprotein, comprising several structural domains, that mediates the specific binding and uptake by receptor-mediated endocytosis of lipoproteins containing either ApoB or ApoE [26]. The protein is synthesized in the rough endoplasmic reticulum (ER) as a precursor protein, followed by a maturation process during the transportation to the Golgi apparatus, including elongation of carbohydrate chains, removal of signal peptides (21 amino acid), and conformational change. The 160 kDa transmembrane receptor is present at the surface of most cell types and mediates the transport of LDL-C into cells, through receptor-mediated endocytosis, thus playing a pivotal role in cholesterol homeostasis [27].
The mature form of the LDL receptor protein contains 839 amino acid residues and five functional domains: the ligand-binding domain, the epidermal growth factor (EGF) precursor homology domain, the *O*-linked polysaccharide domain, and the cytoplasmic domain (Fig.1). Different domains of the LDLR are responsible for different steps during lipoprotein uptake.

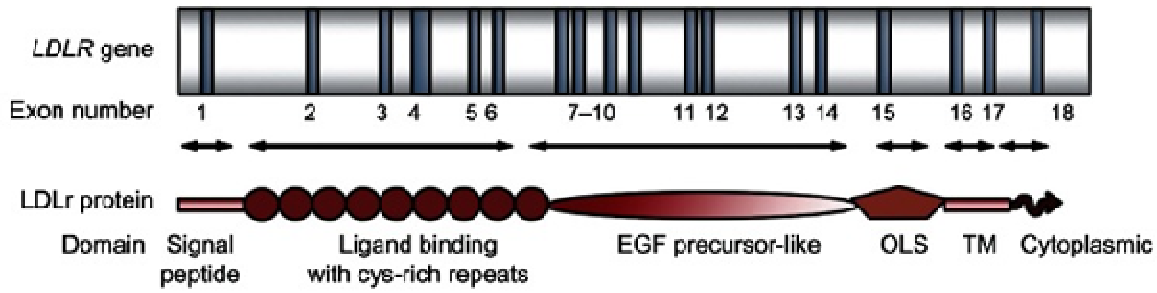

Figure 1. Correlation of exon organization with protein domains in the LDL receptor

The ligand-binding domain contains seven tandem repeats (LA repeats 1-7), each repeat contains six cysteine residues that form three intra repeat disulphide bonds. The LDLR uses LA repeats 4 and 5 to bind to ApoE on chylomicrons, VLDL and LA repeats 3-7 to bind ApoB on LDL-C.
The EGF precursor is 33% identical to a portion of the human epidermal growth factor. This domain is required for the acid-dependent dissociation of lipoproteins from the receptor in the endosome during receptor recycling.
The *O*-linked polysaccharide domain is region building sugars.
The cytoplasmic domain, the membrane-spanning domain that anchors the protein to the cell membrane.
The mutations in the *LDLR* gene in FH patients have helped to delineate the crucial steps of receptor mediated endocytosis. Many different mutations have been identified, and their effect on the structure and function of the protein in cultured cells, has shown that mutations in different regions of the protein can have very different effects on receptor function.
Sequence variations have been categorized into 5 different classes based on biophysical and functional characteristics (Fig.2)

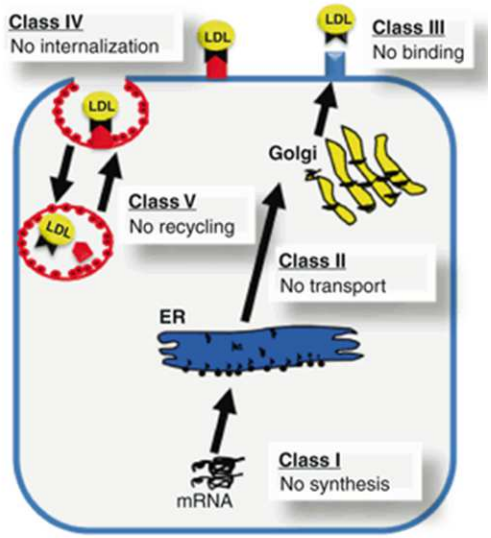

Figure 2. The figure iIllustrates the five different classes for defects of LDL receptor function

Class 1. mutations fail to produce immune-precipitable proteins, no detectable LDLR proteins. This is the most common class of mutant alleles, accounting for approximately half of the mutations so far analyzed.
*Class 2.*mutations encode proteins that are blocked, either completely (Class2A) or partially (Class2B), in transport between the ER and Golgi apparatus, transport defective alleles. These mutant receptors do not appear on the surface of the cells. This is the second most frequent class of mutations.
*Class 3.*mutations encode proteins that are synthesized and transported to the cell surface, but fail to bind LDL normally. Defective LDL binding.
*Class 4.*mutations encode proteins that move to the cell surface and bind LDL normally, but are unable to cluster in clathrin-coated pits and thus do not internalize LDL. Internalization-defective alleles.
*Class 5.*mutations encode receptors that bind and internalize ligand in coated pits, but fail to release the ligand at acidic pH in the endosome and fail to recycle to the cell surface. Recycling-defective receptors [28].
Mutations in the LDLR gene are the most frequent cause of ADH accounting for about 85 to 90% of FH cases [8].
More than 1200 different mutations have been described in the literature so far, but only 79% are recognized to be pathogenic (http://www.ucl.ac.uk/fh) [29].
LDLR mutations include mainly single nucleotide changes, which alter the amino acid composition of the mature protein, affect the correct splicing of the transcript, or binding of key transcription factors, if located in the promoter region. Large deletions and insertions account for approximately 5-6%of all FH genetic defects.
Although LDLR defects are primarily identified by genetic methods, it is extremely important to demonstrate a deficiency in the LDLR function of FH suspect patients through functional studies. Receptor assays reported include measurement of radiolabeled-LDL or fluorescently-labeled LDL binding and/or uptake in skin fibroblasts [25] or leukocytes [30].
The Epstein-Barr virus (EBV)- transformed B-lymphocytes show high LDLR levels and allow an unlimited supply of cells without being influenced by the patient's diet and drug treatment [31]. The main problem of these functional assays was the bad separation between residual LDLR activity from FH patients and controls. The separation of different leukocyte populations obtains more accurate results [32]; the selection of viable lymphocytes by stimulation with a mitogen or by flowcytometry gating improves the discrimination between patients and controls although it does not allow complete discrimination of FH heterozygote patients from healthy controls [33].

In a recent work, Romano et al demonstrated the possibility to discriminate between FH patients carrying *LDLR* mutations and healthy controls by an improved functional assay based on the DiI-LDL uptake in ionomycin plus PMA-stimulated T-lymphocytes and they have also identified a cut-off value with high sensitivity and specificity [34]. Many mutations in the LDLR gene have been described in various countries around the world and disease incidence appears to be higher in certain nationalities and geographical areas with elevated cardiovascular risk [35]. The specific point of mutations, deletions, frameshifts, inheritance patterns and geographical distribution can be reviewed on line through the “Online Mendelian Inheritance in Man” web site (OMIM, Mckusick-Nathans Institute for Genetic Medicine, JohnsHopkinsUniversity (Baltimore, MD) and National Center for Biotecnology Information, National Library of Medicine (Bethesda, MD), 2000. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim/).

The distribution of causative FH mutations in Europe varies between countries; some populations, such as French and Italian, demonstrate considerable heterogeneity, while others involve a narrow range of causative mutations and are relatively homogenous [35]. The mutation spectrum in Italy consists of over 80 mutations including large rearrangements, missense and nonsense mutations. Some of these mutations are specific to certain regions of Italy; for example, the p.Gly549Asp mutation is found predominantly in the southern regions of Italy [36]. Moreover, it seems that some of these genetic changes are present in areas of relatively high incidence (FH-clusters) and may represent a “founder effect”. A founder effect occurs when a subpopulation is formed through the immigration of a small number of "founder" subjects, followed by population expansion. If, by chance, some of the founders had FH, then genetic drift could lead to a high proportion of affected subjects who share specific mutations introduced by the founders.

The high number of different FH mutations makes genetic testing labour-intensive and costly, which has encouraged the development of novel assays and techniques such as next-generation sequencing (NGS) for diseases like FH [37].

1.3.2 Apolipoprotein-B gene

In humans, most of plasma cholesterol is transported in LDL. This lipoprotein consists of a central core composed of apolar neutral lipids (triglycerides and cholesteryl esters) and an outer shell of phospholipids, unsterilized cholesterol, and a single ApoB molecule its protein component (Fig.3) [38]. ApoB plays a central role in intracellular assembly and secretion of triglyceride-rich (TG-rich) lipoproteins and is a ligand for LDLR [39].

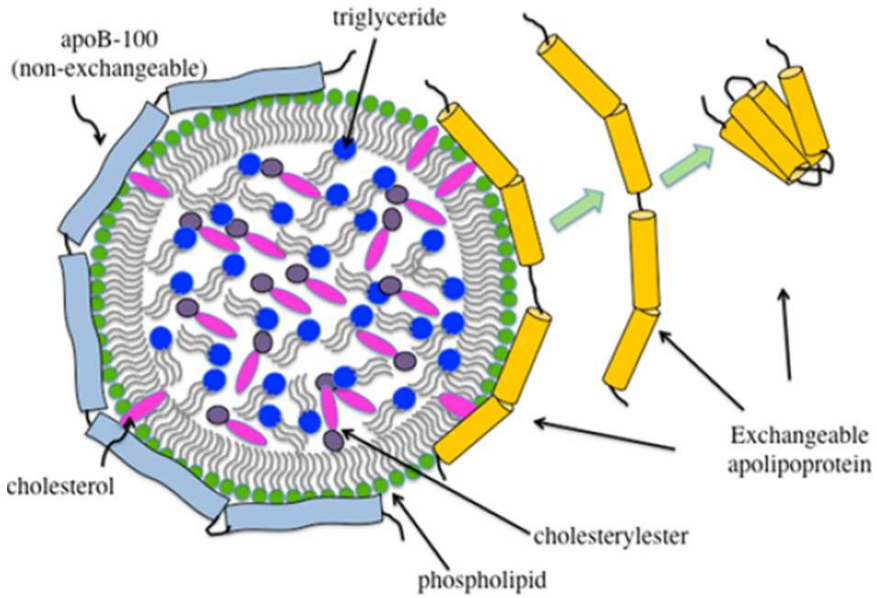

Figure 3. Schematic representation of a generic lipoprotein particle.

The ApoB gene is 43kb in length, contains 28 introns and 29 exons, and is located on the short arm of chromosome 2 and mRNA is 14.5kb in length and codes for a mature protein of 4536 amino acids. The protein has four well-recognized functional actives. First, it is required for the synthesis, assembly, and secretion of hepatic TG-rich lipoproteins. Second, it binds to lipids and is a structural component of VLDL and LDL. Third, it binds to heparin and various proteoglycans found in the arterial wall. Finally, ApoB mediates the uptake of LDL particles by liver and peripheral tissue via a specific interaction with the LDLR [2].

ApoB is the ligand for the LDLR and sole apolipoprotein on the surface of LDL; it is wrapped around the LDL particle as a belt and there are some critical residues that are crucial for the ApoB-LDLR affinity within this configuration [40]. The LDLR-binding domain was localized to the region between residues 3386 and 3396, it was shown that alterations involving this amino acid destabilize important clusters for the APOB conformation, altering in this way its affinity for the LDLR; additionally, regions between amino acid 3174 and 3184 and also 4181 and 4540 were proved to be important for the correct folding of the carboxyl terminus of APOB necessary for binding to the LDLR [41].

Goldestein was the first to indicate that a defect in apoB-100 can lead to a phenotype similar to FH by observation of decreased fractional catabolic rates of autologous LDL compared with homologous LDL in five subjects [42]. *APOB* mutations causing FH are estimated to occur in 1:300 to 1:700 people in several Caucasian populations in Europe [35]. The penetrance of APOB mutations has been shown to be <100% and patients with *APOB* mutations usually have a less severe phenotype than FH patients due to LDLR mutations [43]. In contrast to the large number of variations identified in the LDLR gene, only a few variations have been characterized in the ApoB gene. The most frequent mutation is p.Arg3500Gln that causes defective binding of the LDLR with is ApoB ligand, leading to decreased clearance of circulating LDL. The Arg3500Gln mutation in *APOB* is the most common cause of hypercholesterolemia due to an *APOB* mutation, accounting for 5 to 10% of FH cases in Northern European populations (rare in other populations) [8].

The binding affinities of p.Arg3500Gln to the LDL receptor is reduced to 30% [44]. Interestingly, this mutation is not located at the LDLR-binding site. Instead, an Arg3500-Try4369 interaction is necessary to ensure the proper conformational shape of the ApoB protein, and mutations in these key amino acid result in improper protein folding and reduced receptor binding [41]. However, a large number of ApoB gene mutations truncating ApoB have been reported to be cause of Familial Hypobetalipoproteinemia (FHBL) [45]. Missense mutations of the *APOB* gene can also be the cause of FHBL and are associated with a decreased secretion of the mutant APOB because of an increased binding to microsomal triglyceride transfer protein[46]. Mutations in *APOB* that reduce translation or secretion, or enhance the catabolism of *APOB* are associated with reduced plasma LDL-C levels and reduce risk of coronary disease [47]. In line with this, patients with extremely low levels of APOB (< 5th percentile) due to FHBL seem to be protected against CVD [48]. Based on these observations APOB is conceptually an attractive target to reduce CVD risk.

1.3.3 Proprotein convertase subtilisin/kexin type 9 gene

PCSK9 is the ninth member of the Proprotein Convertase (PC) family that has profound effects on plasma LDL-C levels through its ability to mediate LDL receptor protein degradation. The *PCSK9* gene is localized on chromosome 1p32.3. This gene is about 22kb and comprises 12 exons encoding a 692 amino acid glycoprotein. This convertase is highly expressed in the liver, small intestine, and brain, kidney and is present in human plasma [49].
PCSK9 is synthesized as a 74 kDa soluble zymogen (proPCSK9) that undergoes autocatalytic cleavage at position 152 in the endoplasmic reticulum to release the propeptide (14 kDa) from the N-terminus resulting in a mature enzyme of about 60 kDa [50] (Fig.4). The cleavage of the pro-domain is strictly required for PCSK9 maturation and activation.

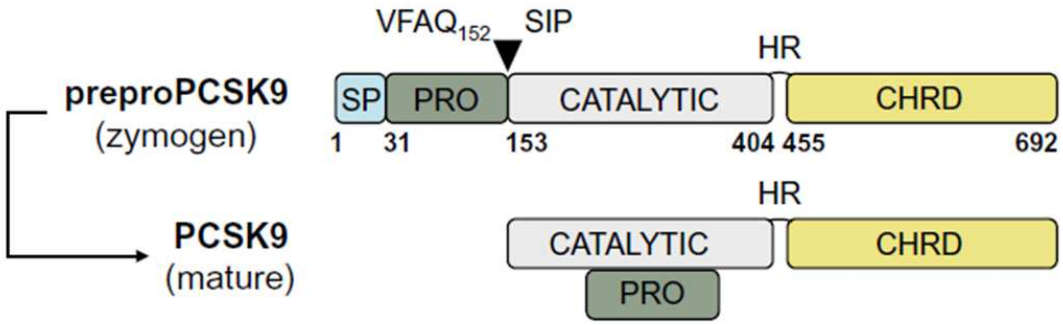

Figure 4. Proprotein convertase subtilisin/kexin type 9 (PCSK9) structure

The protein domains that comprise PCSK9 are a pro-peptide or inhibitory pro-domain, the subtilisin-like catalytic domain, and a cysteine-rich, unique, C terminal domain amino acids[51](Piper DE et al Structure 2007). PCSK9 complexes with the LDLR and is taken up together in clathrin coated pits by hepatocytes. The region within the LDLR that is bound by catalytic domain of PCSK9 corresponds to EGF-A located within the EGF receptor precursor homology domain [52].
PCSK9 binding to this site is absolutely required for LDLR degradation, even though other regions of the LDLR that do not interact with PCSK9 are important for PCSK9-mediated LDLR degradation [53].
A recent study has demonstrated that a soluble EGF-A peptide acts as decoy and inhibits the binding of PCSK9 to LDLR thus providing evidence

for the crucial role of the EGF-A domain in the PCSK9 regulatory pathway [54]. Secreted PCSK9 in the media of cultured cells or in human plasma binds to the EGF-A of the cell-surface LDLR. PCSK9 bound to the LDLR is subsequently internalized by endocytosis, the PCSK9-LDLR complex is located in the sorting endosomes. In contrast to the binding of LDL to the LDLR, PCSK9-LDL receptor affinity is increased at an acidic endosome pH and failure to release PCSK9 prevents receptor recycling and directs the PCSK9-LDLR complex to the lysosome, where degradation of the LDLR occurs.PCSK9 also targets the LDLR within the cell independently of endocytosis. In fact, adenovirally mediated PCSK9 overexpression accelerates the degradation of the LDLR in a post endoplasmic reticulum compartment [55] (Fig.5).
PCSK9 thus regulates the number of LDLR on the cell surface of hepatocytes. Findings also demonstrate that the LDL receptor plays a critical role in facilitating the trafficking of PCSK9 from the endoplasmic reticulum to downstream sites in the secretory and endocytic pathways.

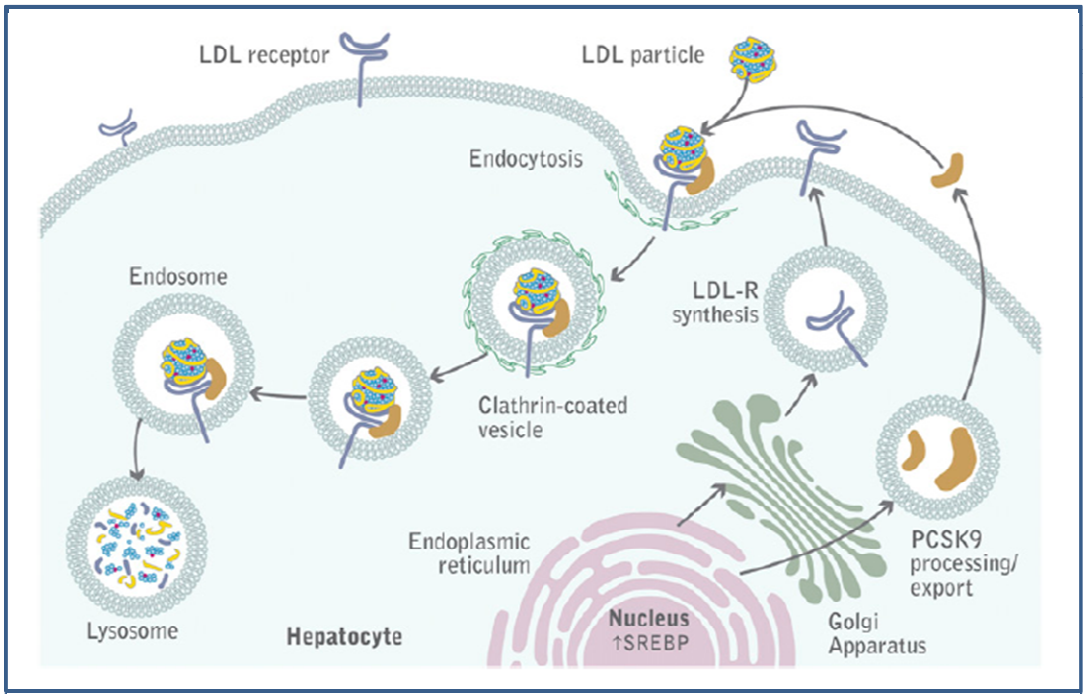

Figure 5. PCSK9-mediated degradation of the LDLR. A complex of LDL-C, LDLR and PCSK9 is internalized and undergoes lysosomal degradation.

Mutations in *PCSK9* have been found to decrease or increase the ability of PCSK9 to degrade the LDLR. Loss-of-function mutations which reduce the ability of the mutant PCSK9 to degrade the LDLR, result in a higher number of cell-surface LDLR and reduced levels of LCL-C in plasma [56]. Conversely, gain-of-function mutations which increase the ability of the mutant PCSK9 to degrade the LDLR, result in a lower number of cell-surface LDLR and increased levels of LDL-C in plasma [56]. Thus, mutations in PCSK9 would be associated with the development of hypercholesterolemia or hypocholesterolemia, respectively. Gain-of-function mutations in PCSK9 cause fewer than 5% of hypercholesterolemia cases in most studies.

Phenotypic variations in patients heterozygous for a mutation in the PCSK9 gene has not yet been evaluated. The clinical comparison of PCSK9 mutation carriers to subjects with mutations in the LDLR or ApoB is difficult to estimate to date as most PCSK9 mutation carriers were identified in a sample selected for clinical diagnosis of FH [3]. After the initial report that gain-of-function mutations in the PCSK9 gene cause hypercholesterolemia, PCSK9 has gained large interest as a target for lipid lowering.

PCSK9 has been shown to be a pivotal regulator of LDL-C metabolism thanks to its role in lysosomal degradation of LDL receptor within hepatocytes. The notion that loss-of-function PCSK9 mutations confer an 80% CVD risk reduction has further substantiated the role of PCSK9 as a potential target for therapy [57].

1.4 *Autosomal Recessive hypercholesterolemia (ARH) and LDLRAP1 gene*

In 1973, Khachadurian and Uthman first described what is now termed Autosomal Recessive hypercholesterolemia (ARH) [58].
ARH is caused by mutations in the LDL receptor adaptor protein 1 gene, which encodes an adaptor protein required for normal LDL receptor-mediated endocytosis in hepatocytes and governs the clustering of the LDLR into clathrin-coated pits [58].
The gene spans about 25 kb and has nine exons and eight introns. The predicted amino acid sequence contains a 170–amino acid motif that shares considerable sequence similarity with the phosphotyrosine binding (PTB) domains of many adaptor proteins. PTB domain (residues 40-180) binds the consensus sequence NPXY, which is present in the cytoplasmic domains of several cell-surface receptors, including the epidermal growth factor receptor, the insulin receptor, nerve growth factor receptor (TrkA) and the LDLR [59, 60].
ARH is a rare disorder in most countries with the exception of the island of Sardinia(Italy), where the disease is not uncommon (frequency1:40,000 for homozygotes and compound heterozygotes), probably as the result of a founder effect and inbreeding [61]. ARH has been considered a phenocopy of homozygous FH due to LDLR mutations. Plasma levels of LDL-C in ARH patients tend to be intermediate between those in FH heterozygotes and those in FH homozygotes. However, the prevalence for developing early onset CHD is lower compared to FH patients with strong homozygous LDLR mutations [62] but patients with ARH often have large, bulky xanthomas [63].

1.5 *Polygenic basis for hypercholesterolemia*

Recently, the European Atherosclerosis Society (EAS) have also released guidelines for the management of dyslipidaemias. They reported that some 10%-40%, depending on referral criteria, of patients with a clinical diagnosis will not have a detectable causal mutation; rather, they have a clinical diagnosis of FH but not a mutation diagnosis. There may therefore be yet genes implicated in this disease; alternatively, these individuals may present a polygenic basis for their LDL cholesterol elevation without contributions from any of the classical FH genes (Fig.6).

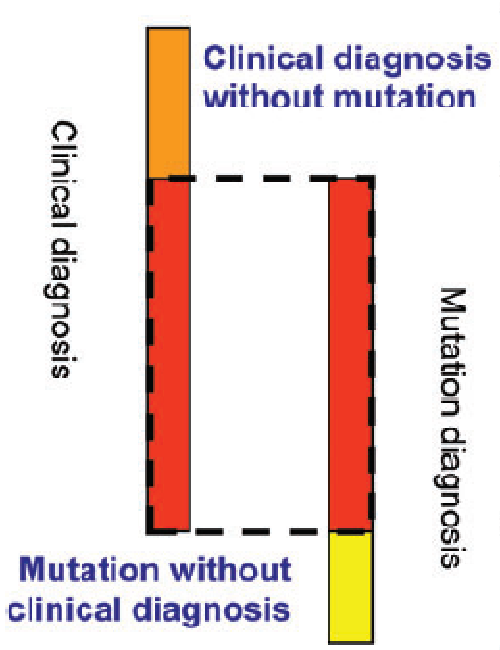

Figure 6. Overlap of clinical and mutation diagnosis of heterozygous FH

Recent genome-wide association studies (GWASs) have localized common DNA sequence variants that contribute to many human phenotypes. The success of this approach has been particularly notable for blood lipoprotein levels.
Small-scale studies have also shown that some individuals carry multiple mutations in different genes, with established roles in lipid metabolism in families with apparent Mendelian forms of dyslipidemia [64].

Since the beginning of large-scale GWAS in 2008 until today, the number of candidate genes associated with plasma lipid phenotypes has steadily increased [65]. In fact, the Global Lipid Genetic Consortium (GLGC) meta-analysis of GWASs studies identified several loci where common variants affect LDL-C concentration, and results of another study showed that a proportion of individuals carrying several LDL-raising single nucleotide polymorphisms (SNPs) have elevated levels of LDL-C [66].
A recent study in English and Belgian FH patients [67], has reported that in a proportion of patients where no mutation can be found in these genes, the disease could be polygenic, due to SNPs strongly associated with LDL-C. However, a proportion of all clinically diagnosed cases of FH could be polygenic, due to the inheritance of a greater than average number of common LDL-C raising alleles leading to an increase in LDL-C above the diagnostic cutoff.

1.6 *Treatment for FH*

The aim of treatment of FH patients is the reduction of plasma LDL-C levels to decrease the risk of premature CVD. All subjects with FH and their families should undergo intensive education targeting lifestyle management, including a healthy diet, exercise, weight control, blood pressure control, and smoking cessation [68].
The most popular and commonly prescribed cholesterol-lowering agents are statins, the 3β-hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR) inhibitors. HMGCR catalyzes the rate limiting step in the biosynthesis of sterols, cholesterol and isoprenoids. The statin class of cholesterol-lowering drugs have been used for decades to successfully lower plasma cholesterol concentrations and cardiovascular risk.
Statin drug therapy significantly reduces the morbidity and mortality from premature CHD in heterozygous FH patients lowering cholesterol and stimulating the production of the LDLR in livers cells [9]. Clinical endpoint trials of lipid-lowering drug therapy with statins have demonstrated their effectiveness in the primary and secondary prevention of coronary heart disease risk, especially in the highest risk groups [69].
Although there are no randomized clinical trials designed specifically for patients with familial hypercholesterolemia, observational studies strongly suggest that statins reduce disease risk in FH individuals [70]. Although in some heterozygous FH and polygenic hypercholesterolemia individuals, LDL-C treatment goals cannot be reached with the maximum available or tolerated dose of statin, these patients might be administered a statin therapy in combination with other lipid-lowering therapies including ezetimibe and resins [71].
Patients with homozygous FH generally respond to a lesser degree compared with heterozygous FH individuals. In patients with homozygous FH who are inadequately controlled on statins alone, the current practice is to perform LDL apheresis at weekly intervals to achieve the lowest possible LDL cholesterol levels [72]. It is generally agreed that targets for LDL-C treatment in children need not be as low as those of adults.
Like in adults good evidence for an absolute or relative target does not exist in children. Numerous clinical trials of the safety and efficacy of statins for lowering LDL cholesterol have been conducted in children and adolescents with severe dyslipidemia. Current pediatric guidelines recommend to start the statin therapy at the age of 8 years or older. The drug therapy in children aged under eight is only recommended in cases of marked elevation of LDL-C usually found in homozygous forms of FH [73] [22].

1.7 *New treatments for FH*

Current therapeutic approaches with lipid-lowering agents, such as statins, fail to protect more than half of patients from cardiovascular events. High statin doses are often associated with an increased frequency of adverse effects. Therefore, there is a need for additional and alternative treatment options. There are several promising novel therapeutic approaches for the treatment of dyslipidaemia and atherosclerosis, which are expected to be of great benefit for patients at risk of cardiovascular disease.
Among the emerging therapeutic options, the most advanced one is the use of antisense oligonucleotides (ASOs) to ApoB-100 mRNA.
Mipomersen or ISIS 301012 is a second-generation antisense oligonucleotide consisting of 20 base pairs, complementary to the coding region for human ApoB-100 mRNA [9]. Mipomersen has been shown to reduce serum LDL-C in patients with homozygous, heterozygous FH and severe hypercholesterolemia [74]. Mipomersen can induce hepatic staetosis in 16% of patients and elevations in plasma amino transferases in 8% of patients [75]. Microsomal Triglyceride transfer protein (MTP) plays an important role in the formation of ApoB. Consequently, MTP inhibition could lead to reductions in the secretion of atherogenic lipoproteins from the liver and also reduce the assembly and secretion of chylomicrons in the intestine [76]. However, in the US Mipomersen (has orphan drug status and, because of the risk of hepatotoxicity), can only be prescribed through a Risk Evaluation Mitigation Strategy (REMS) programme as it is an orphan drug with a risk of hepatotoxicity.
The target for a new therapy is Microsomal triglyceride transfer (MTP) which localizes to the endoplasmatic reticulum of hepatocytes and enterocytes. It transfers triglycerides to the nascent APOB to produce VLDL and chylomicrons. Lomitapide is an oral MTP inhibitor that is intended to treat homozygotes. However, pharmacological inhibition of MTP has resulted in significant increases in hepatic steatosis [11].
Another possible target of therapy is PCSK9. A number of strategies to specifically lower PCSK9 activity are currently in different stages of development and testing. The most advanced strategy is the use of small interfering RNAs and monoclonal antibodies against PCSK9: a monoclonal antibody can inhibit PCSK9 binding to LDLR attenuating PCSK9-mediated reduction in LDLR protein concentration [77].
The inhibition of PCSK9 is the most attractive new approach to reducing atherogenic lipoproteins and enhancing the efficacy of statins. Current data suggest that these drugs will provide an effective therapeutic option for LDL-C reduction and**,** if proven safe in phase III trials, they will be equally

important to LDL-C control [78]. However, Mipomersen and Lomitapide are approved specially for homozygotes FH.PCSK9 inhibitors are not yet licensed for use in FH, but clinical trial data suggest that they may have broad applications for patients with heterozygous FH who do not respond to or who are intolerant of statins [79].

2. Aims

Familial Hypercholesterolemia is an inherited genetic defect resulting in elevated serum LDL-cholesterol concentrations.
The clinical phenotype of FH is known to be due to mutations in three genes encoding proteins involved in the clearance of LDL-C from plasma, LDLR, APOB and PCSK9.
Genetic screening for FH is generally useful to confirm the diagnosis in FH patients and it is essential if the clinical diagnosis is uncertain. The Identification of a causative mutation may provide an additional motivation to implement an appropriate treatment for some patients.
The aims of this study are:
To identify new and known mutations in the three candidate genes, in an attempt to determine to what extent the type of genetic defects is an important determinant of the FH phenotype.
To assess the frequency and spectrum of mutations recognised causing FH among patients attending the different Italian lipid clinics.
To investigate genotype-phenotype correlations in FH carrying different mutations of the LDL-R gene in order to show how the type of the LDLR mutation influences the lipid profile.
To evaluate which of diagnostic criteria DCLN or Simon Broome is more accurate to detect patients with mutations in our population.
In addition, our aim is also to identify the presence of a likely polygenic cause due to the inheritance of LDL-C-raising SNPs which increases LDL-C concentration in a proportion of patients with and without mutations in LDLR, APOB, PCSK9.

3. Materials and Methods

3.1 *Studied population*

This study includes a total of 432 patients with the clinical diagnosis of FH, of which 322 probands were apparently unrelated. The cohort of patients were enrolled at the different Italian lipid clinics. The FH diagnosis was based on personal and family history, physical examination, and untreated plasma low-density LDL-C level >4.9mmol/l. Patients with hypercholesterolemia due to secondary causes including diabetes, hypothyroidism and nephrotic syndrome were excluded. Premature coronary artery disease and risk factors for heart disease, including diabetes mellitus, hypertension, family history of premature coronary artery disease in first-degree relatives, and smoking habits, were recorded. Family members with identified mutations were also invited to participate in this investigation. The cohort consisted of 214 FH-mutation positive (FH-M^+) and 108 FH-mutation negative (FH-M^-) patients. The control group was made of 150 chromosomes of normolipidemic subjects with no history of FH or premature CHD. All novel variations found were searched in Exome Sequencing project (ESP). The study protocol was approved by the institutional review board of the hospital, and informed consent was obtained from each patient or, in the case of children, from their parents.

3.2 *Biochemical analysis*

Biochemical markers were measured on serum after an overnight fast. Total cholesterol, triglyceride and glucose levels were evaluated by standard methods using an automated analyser (Modular P3, Roche). The HDL cholesterol concentrations were measured using a homogeneous enzymatic colorimetric assay and the automated Modular P3 analyzer (Roche). The LDL cholesterol concentrations were calculated according to the Friedewald formula (Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18:499-502) or measured using a homogeneous enzymatic colorimetric assay if triglycerides levels are greater than 4.52mmol/l. LDL particles separation was performed by Lipoprint System (Quantimetrix Inc.).

3.3 *Mutation screening*

Genomic DNA was extracted from whole blood using a Nucleon BACC Genomic DNA Extraction Kits (GE Healthcare Life Sciences).
A screening of the genes associated with the disease was performed to confirm the diagnosis of FH. The promoter, 18 coding exons, and flanking intron regions of the *LDLR* gene were amplified by polymerase chain reaction (PCR) using True Allelic Premix (Applied Biosystems) according to the manufacturer's instructions.
PCSK9 mutation screening was performed through amplification of the 12 exons with flanking intron sequences, whereas APOB screening included the extended binding region of the APOB gene using True Allelic Premix (Applied Biosystems) according to the manufacturer's instructions.
The amplicons were amplified in 30 μL containing 250 ng DNA and 15 pmol of each primer.
Direct bidirectional sequencing analysis of purified PCR product was carried out using the BigDye terminator cycle sequencing ready reaction kit and an ABI Prism 3100 DNA genetic analyzer (Applied Biosystems). CodonCode Aligner version 4.1. software was used for sequence analysis. When new mutation was detected, another sequencing reaction was performed both on genomic DNA from the relative and from a new PCR product from the proband. The amino acid numbers of the LDLR, APOB and PCSK9 are in agreement with the international nomenclature which starts from the initiating methionine and includes the amino acids of the signal peptide (http://www.hgvs.org/mutnomen/).
The causative role of the new mutation was verified according to the following criteria: 1. the absence of the mutation in 150 chromosomes from normocholesterolemic individuals; 2. the conservation of the substituted amino acid residue across homologous proteins by BLASTP; 3. the segregation of the variant with the FH phenotype in the family.

3.4 *In silico analysis*

The potential clinical impact of new variations in LDLR, APOB and PCSK9 genes was assessed using algorithms on pathogenicity predictive websites: PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2), SIFT Human Protein (http://sift.jcvi.org), Mutation Taster (http://www.mutationtaster.org) and PMut (http://mmb.pcb.ub.es/)
The effect of intronic variants on splicing was performed using NetGene2 (http://www.cbs.dtudk/services/NetGene2), Human Splicing Finder (htt://umd.be/HSF/HSF.html) and Automated Splice Site Analysis

(https://splice.uwo.ca). These programs give predictive scores for splice acceptor and donor sequences for wild type and variant sequences. Alterations in the scores between wild type and variations allow an assessment to be made regarding whether or not normal exon splicing will be affected.

3.5 *Multiplex Ligation-dependent Probe Amplification (MLPA)*

MLPA is a rapid, high-throughput technique for copy number quantification and methylation status analysis of genomic sequences.
The SALSA P062-B *LDLR* MLPA kit was obtained from MRC-Holland (Amsterdam, The Netherlands). The *LDLR* MLPA kit contained 37 sets of probes, 19 of which were *LDLR* specific and the others were control standards. Reactions were carried out in 200-µl tubes in a model 9700 thermocycler (PE-Applied Biosystems, Foster City, California).
Each MLPA probe consists of a pair of oligonucleotides, one synthetic and one M13 derived. Each of the two oligonucleotides contains one of the two sequences recognized by the polymerase chain reaction (PCR) primers. One probe oligonucleotide contains a non-hybridized stuffer sequence of variable length (19–370 nucleotides). The two parts of each probe hybridize to adjacent target sequences and are ligated by a thermostable ligase. Upon ligation, a single molecule is formed that can be amplified during PCR. Following PCR, amplicons are separated and quantified by capillary electrophoresis.
Peaks on each electropherogram were normalized by expressing their heights a fraction of the total height of all control peaks.
Each probe in an MLPA probe set has a unique length and can therefore be easily identified. The procedure was performed according to the manufacturer's instructions (www.mrc-holland.com). To obtain gene dosage quotients, the results were compared with three control samples analysed at the same time using the Coffalyser data analysis tool (MRC-Holland).

3.6 *Global Lipids Genetic Consortium gene score calculation*

The GLGC reported SNPs that were significantly associated with LDL-C, with a p-value cut-off of less than 5×10 [66]. For the purpose of the gene score calculation only the lead SNP from each locus was selected and if a SNP was associated with more than one lipid fraction it was only used in this score calculation if LDL-C was the lead trait (most strongly associated trait) for the SNP.

The TaqMan assay was performed for 12 common SNPs, in 11 genes (two in *APOE*) that the Global Lipid Genetic Consortium (GLGC) reported as significantly associated with LDL-C. The possibility of polygenic hypercholesterolemia was assessed in 199 patients who had a clinical diagnosis of FH with or without an identified familial hypercholesterolemia-causing mutation. Gene scores of a randomly selected samples from the UK Whitehall II (WHII) study (n=3020) were used as a healthy control comparison group.

For each individual, the LDL-C-specific gene scores were calculated using the weighted sum of the risk allele (Table 4). Individuals with gene score above 1.16, which was the top decile cutoff for the WHII subjects, were considered to have polygenic hypercholesterolemia.

Table 4. Global Lipid Genetic Consortium 12-SNP LDL-C gene score calculation

	Chromosome number	Gene	Minor allele	Common allele	GLGC weight for score calculation
rs2479409	1	*PCSK9*	G*	A	0·052
rs629301	1	*CELSR2*	G	T*	0·15
rs1367117	2	*APOB*	A*	G	0·10
rs4299376	2	*ABCG8*	G*	T	0·071
rs1564348	6	*SLC22A1*	C	T*	0·014
rs1800562	6	*HFE*	A	G*	0·057
rs3757354	6	*MYLIP*	T	C*	0·037
rs11220462	11	*ST3GAL4*	A*	G	0·050
rs8017377	14	*NYNRIN*	A*	G	0·029
rs6511720	19	*LDLR*	T	G*	0·18
rs429358	19	*APOE*	C	T	..
rs7412	19	*APOE*	T	C	..
ε2ε2	19	*APOE*	..	..	−0·9
ε2ε3	19	*APOE*	..	..	−0·4
ε2ε4	19	*APOE*	..	..	0·2
ε3ε3	19	*APOE*	..	..	0
ε3ε4	19	*APOE*	..	..	0·1
ε4ε4	19	*APOE*	..	..	0·2

* Risk alleles (LDL-raising).

3.7 *SNP Genotyping Analysis Using TaqMan Assays*

TaqMan 5'-nuclease assay chemistry provides a fast and simple way to get single nucleotide polymorphism (SNP) genotyping results TaqMan assays contain two specific primers targeting the region flanking the SNP site and two TaqMan fluorescent probes with a Minor Groove Binder (MGB). These TaqMan probe and primer sets (assays) uniquely align with the genome to provide unmatched specificity for the allele of interest.
Each probe is labeled with a different fluorophore (VIC or 6-FAM). These reporters are attached covalently to the 5' end of the two probes. Near the 3 'end, there is a non-fluorescent quencher (NFQ) that prevents liberation of the reporter fluorescence if the probe is not degraded. The MGB serves to stabilize the double-stranded structure formed between the target and the

probe. During the PCR reaction, probes that hybridize specifically to DNA fragments are destroyed and the fluorescence of corresponding fluorophore is liberated. The oligonucleotide is designed such that there are complementary regions at each end and a probe sequence located in between. This design allows the probe to take on a hairpin, or stem-loop, structure in its natural, isolated state. Because of the stem-loop structure of the probe, the fluorophore is in close proximity to the quencher, thus preventing the molecule from emitting any fluorescence. The molecule is also engineered such that only the probe sequence is complementary to the genomic DNA that will be used in the assay [80].

If the probe sequence of the molecular beacon encounters its target genomic DNA during

the assay, it will anneal and hybridize. Because of the length of the probe sequence, the hairpin segment of the probe will be denatured in favour of forming a longer, more stable probe-target hybrid. This conformational change permits the fluorophore and quencher to be free of their tight proximity due to the hairpin association, allowing the molecule to fluoresce. If on the other hand, the probe sequence encounters a target sequence with as little as one non-complementary nucleotide, the molecular beacon will preferentially stay in its natural hairpin state and no fluorescence will be observed, as the fluorophore remains quenched.

The unique design of these molecular beacons allows for a simple diagnostic assay to identify SNPs at a given location. If a molecular beacon is designed to match a wild-type allele and another to match a mutant of the allele, the two can be used to identify the genotype of an individual. If only the first probe's fluorophore wavelength is detected during the assay then the individual is homozygous to the wild type. If only the second probe's wavelength is detected then the individual is homozygous to the mutant allele. Finally, if both wavelengths are detected, then both molecular beacons must be hybridizing to their complements and thus the individual must contain both alleles and be heterozygous.

Reactions were carried out in 384 plate (384 Well Clear Optical Reaction Plate), prepared with Biomek 2000 Workstation (Beckman), in ABI PRISM 7900 HT Real-time PCR system.

After PCR amplification and end-point fluorescence detection, genotypes are scored using SDS 2.3 (Sequence Detection System, Applied Biosystems) analysis software, which displays fluorescence data in a X-Y plot format for easy clustering and scoring.

3.8 *Statistical analysis*

Continuous variables were expressed as a mean ± SD (parametric distributions) or median value and 25-75 percentile (non-parametric distributions) and categorical data as the absolute number and percentage. Kolmogorov-Smirnov test was performed to assess the hypothesis of a normal distribution of variables. Differences between groups were assessed by T-test, non-parametric Mann-Whitney test and chi-square test, as appropriate. Genotype and allele frequencies were calculated by allele counting and departure from Hardy-Weinberg expectation was evaluated by chi-square analysis (all SNPs were in equilibrium). Analysis of variance (ANOVA) was used to analyze the differences between group means and their associated procedure. The Kruskal–Wallis one-way analysis of variance by ranks was used in multiple test when the variable was non-parametric distribution. To correct the multiple testing was used by Bonferroni test.

To test the type of association, dominant and recessive models were constructed for the rare allele of each polymorphism and univariate odds ratio (OR) with 95% Confidence Interval (CI) were calculated.

The linear regression was performed of observed baseline LDL-C concentrations with the weighted LDL-C gene scores. The risk ratio was calculated of having a measured LDL-C concentration of higher than 4.9 mmol/l by deciles of the GLGC weighted score in WHII controls.

Statistical analysis was performed with PASW version 18.0 software (SPSS Inc.).

4. *Results*

4.1 Patients characteristics

In the group of 432 patients, 322 unrelated (age 33 ± 18) were screened for FH mutations in three candidate genes, *LDLR, APOB, PCSK9*. The clinical characteristics of the subjects enrolled are reported in Table 5. There was no significant difference in the male/female ratio between the groups of FH-M^+ and FH-M^- but the age was significantly different between the two groups (p=0.001). The patients bearing mutations were younger than those without any mutation. All lipid parameters showed non-parametric distributions, the data was reported as median and 25-75 percentiles.
The median total cholesterol (8.3mmol/l) and LDL-cholesterol (6.5mmol/l) in FH-M^+ were significantly higher than FH-M^- (p<0.001). There was no significant difference in triglyceride levels between the two groups. The clinical hallmarks of FH such as tendon xanthomatas and corneal arcus showed no significant difference between the groups probably due to the presence of missing data.

Table 5. Basic characteristics of FH patients

	Total n=322	**FH-M^+ n=214**	**FH-M^- n =108**	**Sig. M^+ vs M^-**
Men	127/195 (39.4%)	86/128 (40%)	41/67 (38%)	ns
Age (years)	33 ± 18	30.5 ± 17.4	37.8 ± 18.6	p=0.001
Total cholesterol (mmol/L)	7.9 (7.1-9.2)	8.3 (7.4-9.6)	7.4 (6.7-8.1)	p<0.0001
LDL-cholesterol (mmol/L)	5.0 (6.1-7.1)	6.5 (5.5-7.6)	5.2 (4.5-5.8)	p<0.0001
HDL-cholesterol (mmol/L)	1.3 (1.1-1.7)	1.3 (1.1-1.6)	1.4 (1.2-1.8)	p=0.01
no-HDL –Cholesterol (mmol/L)	6.6 (5.5-7.7)	7.0 (5.9-8.1)	5.8 (5.1-6.5)	p<0.0001
Triglycerides (mmol/L)	1.1 (0.8-1.6)	1.0 (0.8-1.5)	1.2 (0.9-1.6)	ns
Carotid thickning n (%)	49 (40.5)	36 (43.0)	13 (40.0)	ns
Tendon xanthomas n (%)	21 (13.9)	18 (16.0)	3 (7.0)	ns
Corneal arcus n (%)	17 (18.5)	14 (21.0)	3 (13.0)	ns
Carotid Atherosclerotic plaques n (%)	20 (16.5)	17 (10.0)	3 (19.0)	ns
Smoker (%)	43 (26.5)	27 (41)	16 (22)	ns

4.2 *Mutations in the FH causative genes*

In our cohort we found mutations in 214/322 unrelated subjects (mutation rate 66.5%).
All patients with *APOB* or *PCSK9* mutations were heterozygotes; whereas among patients with *LDLR* mutations there were 3 homozygotes (two identical mutations, 1%) and 7 compound heterozygotes (with different mutations on each LDLR allele, 2%). The total number of unrelated patients with 2 mutations was 10/322 (3.1%). (Fig.7)
The frequency of homozygotes/compound heterozygotes in the total population was 17/432 (4 homozygotes, 13 compound heterozygotes) 3.9%.

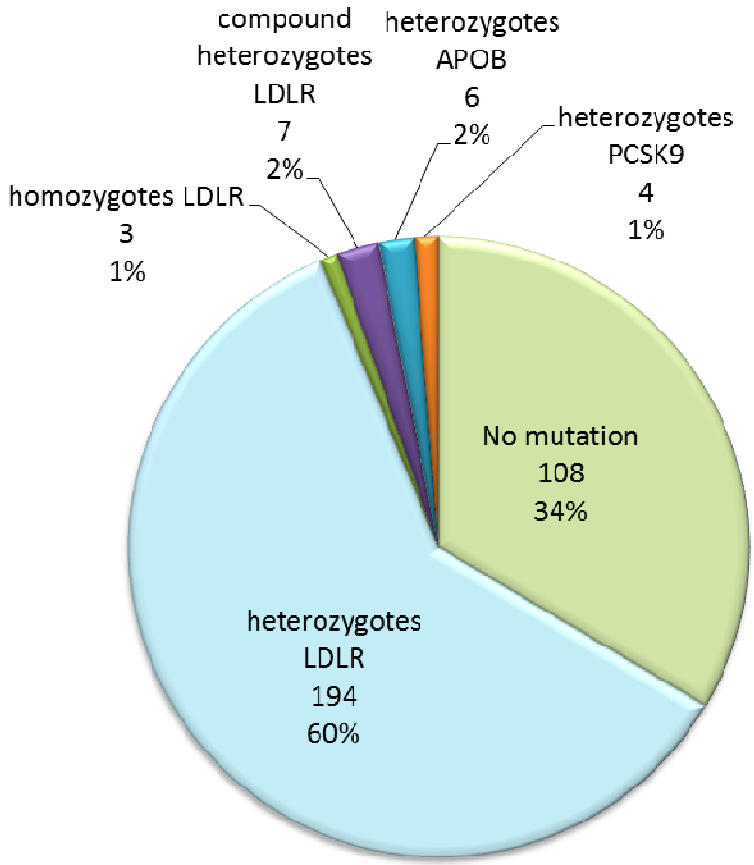

Figure 7. Distribution of mutations in *LDLR, ABPO, PCSK9* genes

4.3 *Spectrum of mutations in LDLR, APOB, PCSK9 genes*

4.3.1 Mutations in the LDLR gene

Among 214 mutated subjects, 204 (95.3%) were carriers of *LDLR* mutations, 6 (2.8%) of *APOB* mutations and 4 (1.8%) of *PCSK9* mutations. All mutations in *APOB* and *PCSK9* were missense mutations.
Among 204 patients with mutations in the LDLR gene we found 214 mutations of which 122 missense (57.0%), 15 nonsense (7.0%), 57 splicing (26.6%), 7 small rearrangements including small deletion, small deletion/insertion (3.3%) and 1 regulatory (0.46%).
Furthermore direct sequencing revealed 58 different mutations, 15 present only in this cohort. We carried out a MLPA analysis to search for large rearrangements and we found gross deletions in 12/214 mutations (5.6%) as shown in figure 8.

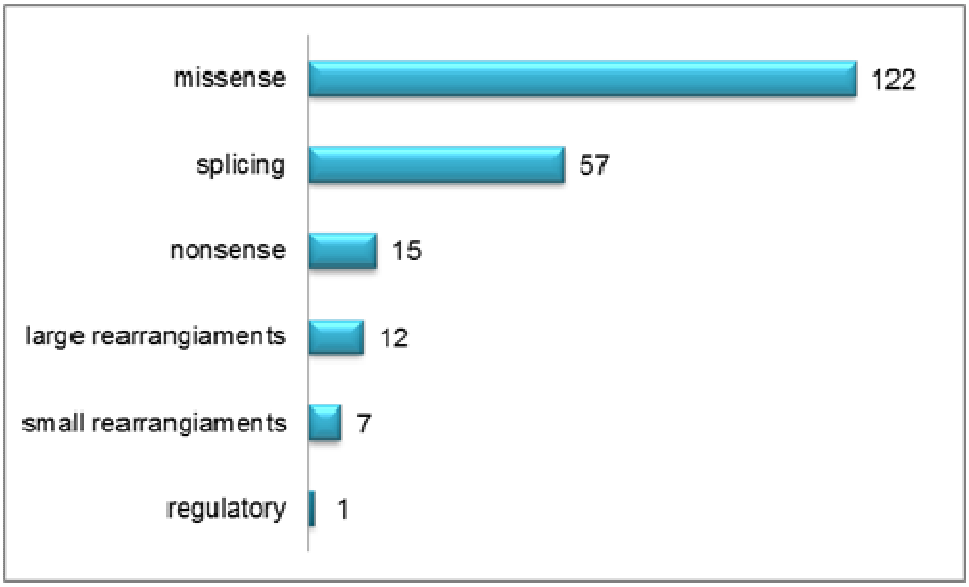

Figure 8. Number of mutation types identified in our cohort

Among the 58 different mutations found, 6 mutations accounted for more than 50% of cases. The most commonly observed mutations in the *LDLR* gene were c.1775G>A (p.Gly592Glu) was found in 31 patients (15.2%), the splicing mutation in intron 15 c.2312-3C>A (causing exon 16 skipping:

p.Ala771_796del) was found in 28 patients (13.7%), the splicing mutation in intron 10 c.1586+1G>A (causing exon 10 skipping: p.Thr454_Gly529del and retention of intron 10: p.Gly529_Phe530ins22) was found in 19 patients (9.3%), the missense mutation c.1567G>A (p.Val523Met) was found in 16 patients (7.8%) and c.1646G>A (p.Gly549Asp) was found in 15 patients (7.3%). In table 5 are reported the 58 different mutations.

Table. 5 *LDLR* mutations identified in our population.

Position	Mutation type	Nucleotide	Protein	Frequency
Promoter	Regulatory	c.1-156 C>T	—	1
Exon 1	Missense	c.58G>A	p.Gly20Arg	1
Exon 2	Nonsense	c.116_117delGCinsAA	p.Cys39X	4
Exon 2	Missense	c.102C>G	p.Cys34Trp	1[N]
Exon 3	Nonsense	c.304C>T	p.Gln102X	5
Intron 3	Splicing	c.313+1G>A	p.Leu64_Pro105delins Ser	1
Exon 4	Missense	c.352G>T	p.Asp118Tyr	2
Exon 4	Missense	c.367T>C	p.Ser123Pro	5[*]
Exon 4	Deletion	c.369_370delTC	p.Ser123fsX6	2
Exon 4	Missense	c.407A>T	p.Asp136Val	2[*]
Exon 4	Deletion	c.424_430delTCCTGCC	p.Ser142ArgfsX62	2
Exon 4	Missense	c.440C>T	p.Thr147Ile	2
Exon 4	Missense	c.465C>A	p.Cys155X	9
Exon 4	Missense	c.514G>A	p.Asp172Asn	1
Exon 4	Missense	c.542C>G	p.Pro181Arg	2
Exon 4	Missense	c.662A>G	p.Asp221Gly	1
Intron 4	Splicing	c.694+1G>C	IVS4+1G>C	1 [N]
Exon 5	Missense	c.761A>C	p.Gln254Pro	1
Exon 6	Missense	c.859G>T	p.Gly287Cys	2
Exon 6	Missense	c.892 A>G	p.Met298Val	1 [N]
Exon 7	Missense	c.953G>T	p.Cys318Phe	1
Exon 7	Missense	c.974G>A	p.Cys325Tyr	2

Position	Mutation type	Nucleotide	Protein	Frequency
Exon 7	Missense	c.1003G>A	p.Gly335Ser	1
Intron 7	Splicing	c.1060+10	p.asp333GlyfsX20	1
Intron 8	Splicing	c.1061-8T>C	IVS7-8 T>C	1
Exon 8	Nonsense	c.1070_1071dupAG p.	p.Cys358SerfsX13	2 [N]
Exon 8	Missense	c.1118G>A	p.Gly373Asp	3
Exon 8	Missense	c.1130G>T	p.Cys377Phe	4 [*]
Exon 8	Missense	c.1135T>C	p.Cys379Arg	16 [*]
Intron 8	Splicing	c.1187-10G>A	p.Gly396AspfsX19	3
Exon 9	Missense	c.1211C>T	p.Thr404Ile	1
Exon 9	Missense	c.1215C>G	p.Asn405Lys	1
Exon 9	Missense	c.1216C>A	p.Arg406Arg	1
Exon 9	Missense	c.1247G>C	p.Arg416Pro	2
Exon 9	Missense	c.1277T>G	p.Leu426Arg	2 [N]
Exon 9	Missense	c.1295T>C	p.Leu432Pro	1
Exon 9	Missense	c.1331C>T	p.Ser444Phe	1
Exon 10	Deletion	c.1478_1479delCT	p.Ser493CysfsX42	3 [*]
Exon 10	Missense	c.1558A>G	p.Arg520Gly	1
Exon 10	Missense	c.1567G>A	p.Val523Met	21 [*]
Intron 10	Splicing	c.1586+1G>A	p.Thr454_Gly529del and p.Gly529_Phe530ins22	25 [*]
Exon 11	Missense	c.1646G>A	p.Gly549Asp	20 [*]
Exon 11	Missense	c.1694G>T	p.Gly565Val	1
Exon 11	Deletion/Insertion	c.1698_1704delinsGCCCAAT	p.Ile566_Leu568 delinsMetProAsn	1
Exon 11-12	Large deletion	c.1587-?_1845+?del	p.Phe530ThrfsX49	10
Exon 12	Missense	c.1730G>C	p.Trp577Ser	3
Exon 12	Missense	c.1739C>T	p.Ser580Phe	6 [*]
Exon 12	Missense	c.1775G>A	p.Gly592Glu	48 [*]

Position	Mutation type	Nucleotide	Protein	Frequency
Exon 13	Missense	c.1898G>A	p.Arg633His	1
Exon 13-15	Large deletion	c.1846-?_2140+?del	p.Asp616ArgfsX16	1
Exon 13-15	Large deletion	c.1846-?_2311+?del	p.Asp616LeufsX17	8
Exon 14	Missense	c.2054C>T	p.Pro685Leu	6*
Exon 15	Splicing	c.2311+1G>A	IVS15+1G>A	1
Exon 15	Nonsense	c.2215C>T	p.Gln739X	1
Intron 15	Splicing	c.2312-3C>A	p.Ala771_Ile796del	47*
Exon 16	Missense	c.892A>G	Met298Val	1
Exon 16	Missense	c. 2389G>A	p.Val797Met	3
Exon 17	Missense	c.2476C>A	p.Pro826Thr	2*

*indicates the presence of a mutation in a compound heterozygotes or in homozygotes

N indicates a novel variation

Variations were detected along the entire *LDLR* gene (Fig.9). To estimate the frequency of mutations by exon compound and multiple mutations we considered them separately. Exon 10 had the highest frequency of mutations (20%), while the frequency in exon 12 and exon 16 was 19.1%-15.1% respectively. Exon 4 (10.3% frequency of mutation) was more susceptible to suffer modifications since it had approximately 17 different mutations, which is the highest number of different mutations reported in exons . The other exons were less susceptible to modifications with a very low mutation frequency.

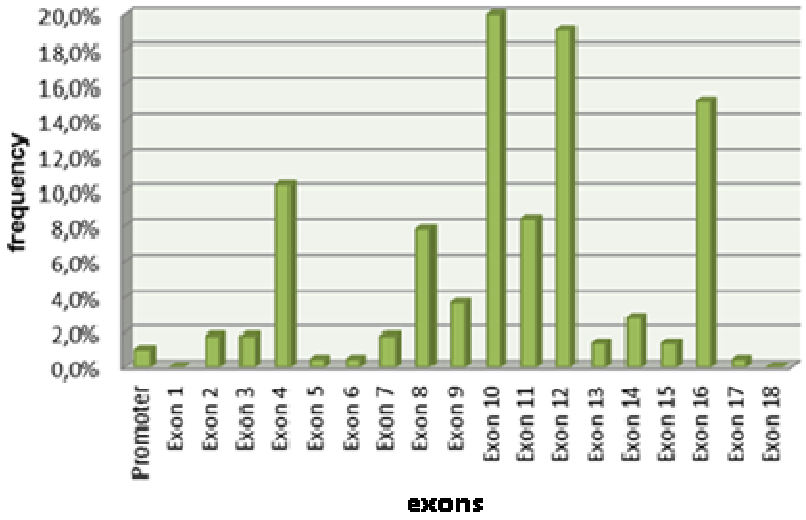

Figure 9. Frequency of point mutations distributed along the *LDLR* gene

In this study were found 5 novel variations which were not reported in the FH database. Three new variations were a single nucleotide change (c.102C>A; c.892A>G; c.1277T>G) resulting in aminoacid substitution p.Cys34Trp; p.Met298Val; p.Leu444Arg respectively. These missense variations affect aminoacid residues conserved across the homologous LDL receptor proteins from 12 different examined species (chimpanzee, gorilla, orangutan, cat, rat, mouse, guinea pig, macaque, dog, pig, cow and zebrafish).

One involved in intron 4 splice donor site (c.694+1G>C), reverse transcription and PCR amplification ware performed for variant c.694+1G>C. The band corresponding to the skipping of exon 4 was more intense than that observed in our control, thereby confirming the reduced use of the donor splice site. We also observed three bands not detectable in the healthy control.

One small duplication of AG in exon 8 (c.1070_1071dupAG) predicting frame shift and premature termination.

4.3.2 Mutations in the APOB gene

In the *APOB* gene group, direct sequencing of exons 26-29 revealed four different mutations (c.9916G>A, c.10672C>T c.10897T>C c.12536C>G) resulting in aminoacid substitution p.Val3306Ile p.Arg3558Cys p.Trp3633Arg p.Thr4179Ser respectively (Table 6).
Two of them p.Val3306Ile and p.Trp3633Arg were novel variations. These mutations affect aminoacid residues conserved across the homologous APOB proteins and were searched in Exome Sequencing Project (ESP), none of which was included in ESP.

Table 6. Mutations identified in the *APOB* gene

Position	Mutation type	Protein	Frequency
26	Missense	c.9916G>A p.Val3306Ile*	1
26	Missense	c.10672C>T p.Arg3558Cys	1
29	Missense	c.10897T>C p.Trp3633Arg*	1
29	Missense	c.12536C>G p.Thr4179Ser	3

* Novel variation

4.3.3 Mutations in the PCSK9 gene

In the *PCSK9* gene group, direct sequencing of the 12 encoding region revealed four different mutations (c.991C>G c.1069C>T c.1394C>T c.1496G>A) resulting in aminoacid substitution p.Pro331Ala, p.Arg357Cys p.Ser465Leu, p.Arg499Hys respectively (Table 7).
Two were known mutations p.Arg357Cys, p.Ser465Leu and other two were novels p.Pro331Ala p.Arg499Hys.
These mutations affect aminoacid residues conserved across the homologous PCSK9 proteins and were searched in Exome Sequencing Project (ESP), none of which was included in ESP.
The functional characterization of p.Ser465Leu was performed as described in our previously paper [81].

Table 7. Mutations identified in *PCSK9* gene

Position	Mutation type	Protein	Frequency
6	Missense	c.991C>G p.Pro331Ala*	1
7	Missense	c.1069C>T p.Arg357Cys	1
9	Missense	c.1394C>T p.Ser465Leu	1
9	Missense	c.1496G>A p.Arg499Hys*	1

* Novel variation

4.5 *In silico analysis of the LDLR, APOB, PCSK9 mutations*

In silico analysis of the novel mutations of the three genes is shown in table 8.

In the case of *LDLR,* only the missense mutation p.Cys34Trp resulted to be pathogenic for four algorithmics. The mutation p.Leu444Arg resulted to be pathogenic based on the consensus of PolyPhen2, SIFT, Mutation Taster, resulted neutral for PMut algorithms. The splicing mutation c.694+1G>C was predicted to abolish the donor splice site by all the 5 algorithms used. *In silico* analysis was not applicable for duplication of AG in exon 8 and it was assessed using conservation and structure analysis.

The causative role of mutations in the *APOB* and *PCSK9* genes resulted uncertain based on the consensus of four algorithms, were considered to be possibly damaging on the basis of indirect criteria such as the aminoacid conservation across the homologous proteins and the segregation with the FH phenotype.

Table 8. *In silico* prediction of the effect of variants in *LDLR, APOB, PCSK9* genes

Gene	Position	Variant	Protein Prediction				Splicing Prediction				
			PolyPhen2	SIFT	Mutation taster	PMut	Splice Site Finder-Like	Max Ent Scan	NNSPLICE	Gene Splicer	Human Splicing Finder
LDLR	Exon 2	c.102C>A (p.Cys34Trp)	D	D	D	D	-	-	-	-	-
	Intron 4	c.694+1G>C IVS4+1	-	-	-	-	Site lost in c-allele	Site lost in c-allele	Site lost in c-allele	Site lost in c-allele	Site lost in c-allele
	Exon 6	c.892A>G (p.Met298Val)	B	B	B	B	-	-	-	-	-
	Exon 8	c.1070_1071dup AG (p.Cys358 Ser fsx13)	N/A	N/A	N/A	N/A	-	-	-	-	-
	Exon 9	c.1277T>G (p.Leu444Arg)	D	D	D	B	-	-	-	-	-
APOB	Exon 26	c.10672C>T (p.Arg3558Cys)	P	D	B	D	-	-	-	-	-
	Exon 29	c.10897T>C (p.Trp3633Arg)	P	D	D	D	-	-	-	-	-
PCSK9	Exon 6	c.991C>G (p.Pro331Ala)	P	B	D	B	-	-	-	-	-
	Exon 9	c.1496G>A (p.Arg499Hys)	B	B	D	B	-	-	-	-	-

D–probably damaging (PolyPhen2), damaging (SIFT), disease causing (Mutation Taster), pathological (PMut)

P–possibly damaging

B–benign(PolyPhen2), tolerated (SIFT), polymorphism (Mutation Taster), neutral (PMut) N/A–not applicable

4.6 *Correlation between lipid levels, mutation type and clinical criteria*

4.6.1 Differences in lipid levels between mutation type/number of mutations

To evaluate the influence of the *LDLR* mutations on the phenotype, the patients were divided into four groups: no mutation, carrier of missense mutation, carrier of one radical mutation, homozygotes/compound heterozygotes.
The missense carrier group included only the single nucleotide change resulting in aminoacid substitution, whereas the radical mutation group included carriers of small deletions (with frame shift effect), large deletions (deletion of entire exons), splicing mutations (causing exon skipping or intron retention), nonsense mutations (giving rise to a premature stop codon producing a truncated protein.
Levels of TC and LDL-C gradually increased in the four group of patients. All of these differences in TC and LDL-C were statistically significant at Kruskal Wallis test ($p<0.0001$).
Multiple testing by Bonferroni correction revealed statistical differences between each group versus the others (Table 9).
There was no significant difference in TG and HCL-C levels between the four groups.

Table 9. Differences between lipid levels and mutations type

	no mutation n=108	Significant at Bonferroni	Missense n=117	Significant at Bonferroni	Radical n=87	Significant at Bonferroni	Homozygotes /Compound heterozygotesn=10	Significant at Kruskal Wallis
TC (mmol/l)	7.4 (6.7-8.1)	Sig vs Miss p=0.021 Sig vs Rad p<0.0001 Sig vs Comp p<0.0001	8.0 (7.0-9.2)	Sig vs Rad p=0.006 Sig vs Comp p<0.0001	8.7 (7.7-10.1)	Sig vs Comp p<0.0001	12.0 (8.7-14.5)	p<0.0001
LDL-C (mmol/l)	5.2 (4.5-5.8)	Sig vs Miss p<0.0001 Sig vs Rad p<0.0001 Sig vs Comp p<0.0001	6.3 (5.0-7.1)	Sig vs Rad p=0.003 Sig vs Comp p<0.0001	6.7 (6.1-7.8)	Sig vs Comp p<0.0001	9.8 (6.7-13.0)	p<0.0001
HDL-C (mmol/l)	1.4 (1.2-1.8)	Sig vs Miss p=0.059 Sig vs Rad p=0.023 Sig vs Comp p=0.011	1.3 (1.1-1.6)	ns vs Rad ns vs Comp	1.3 (1.1-1.6)	ns vs Comp	1.1 (0.7-1.2)	p=0.008
non-HDL-C (mmmol)	5.8 (5.1-6.5)	Sig vs Miss p=0.001 Sig vs Rad p<0.0001 Sig vs Comp p<0.0001	6.6 (5.6-7.7)	ns vs Rad ns vs Comp	7.2 (6.3-8.4)	Sig vs Comp p<0.0001	11.0 (7.6-13.4)	p<0.0001
TG (mmol/l)	1.2 (0.9-1.6)	ns vs Miss ns vs Rad ns vs Comp	1.0 (0.8-1.4)	ns vs Rad ns vs Comp	1.0 (0.8-1.7)	ns vs Comp	1.0 (0.8-1.4)	ns

Miss= missense, Rad= radical, Comp= Homozygotes/Compound heterozygotes
ns= not significant

4.6.2 Differences in lipid levels between groups based on DCLN diagnosis

The DCLN criteria cannot be used in children, for this reason an unrelated paediatric cohort was removed from the analysis and therefore DCLN score was available only for 254 patients.
According to DCLN criteria, patients were divided into four groups: definite-FH (subjects with score >8 points), probable-FH (subjects with scores 6-8 points), possible-FH (subjects with scores 3 to 5 points) and no FH (subjects with scores 0-2 points).
Levels of TC and LDL-C gradually increased from patients without mutations to patients with a possible-FH, to patients with a probable-FH and to patients with a definite-FH. All these differences in TC and LDL-C were statistically significant at Kruskal Wallis test ($p<0.0001$).
Multiple testing by Bonferroni correction revealed statistical differences between each group versus the others (Table 10).
There was no significant difference in TG and HCL-C levels between the four groups.

.

Table 10. Differences in lipid levels among groups based on DCLN diagnosis

	no-FH n=73	Significant at Bonferroni test	possible-FH n=124	Significant at Bonferroni test	probable FH n=32	Significant at Bonferroni test	Definite-FH n=25	Significant at Kruskal Wallis
TC (mmol/L)	6.7 (5.8-7.8)	Sig vs pos p<0.0001 Sig vs pro p<0.0001 Sig vs def p<0.0001	8.0 (7.4-8.9)	Sig vs pro p=0.004 Sig vs def p<0.0001	9.2 (8.0-10.3)	Sig vs def p=0.030	10.1 (9.1-12.6)	p<0.0001
LDL-C (mmol/L)	4.4 (4.0-4.8)	Sig vs pos p<0.0001 Sig vs pro p<0.0001 Sig vs def p<0.0001	6.0 (5.4-6.7)	Sig vs pro p<0.0001 Sig vs def p<0.0001	7.2 (6.6-8.5)	Sig vs def p=0.001	8.1 (7.3-10.4)	p<0.0001
HDL-Cl (mmol/L)	1.4 (1.2-1.8)	ns vs pos Sig vs pro p=0.030 ns vs def	1.3 (1.1-1.7)	ns vs pro ns vs def	1.3 (1.1-1.4)	ns vs def	1.3 (1.0-1.6)	ns
Cholesteron -no-HDL (mmol/L)	4.9 (4.6-5.2)	Sig vs pos p<0.0001 Sig vs pro p<0.0001 Sig vs def p<0.0001	6.6 (6.0-7.2)	Sig vs pro p<0.0001 Sig vs def p<0.0001	7.9 (6.8-9.1)	Sig vs def p=0.002	8.5 (7.7-11.0)	p<0.0001
TG (mmol/L)	1.2 (0.7-1.3)	ns vs pos ns vs pro ns vs def	1.1 (0.9-1.7)	ns vs pro ns vs def	1.3 (0.9-1.6)	ns vs def	1.2 (0.8-1.6)	ns

4.6.3 Lipid levels differences among groups based on the Simon Broome diagnosis

Since the Simon Broome diagnostic criteria for FH include pediatric subjects a pediatric cohort was also included in our analysis . According to these criteria our cohort was divided into three groups: possible-FH, probable-FH and no-FH. In this last group we included subjects with TC and LDL-C levels above the Simon Broome criteria cut off (7.5mmol/l and/or 4.9mmmol/l respectively) but without family history that could be clarified.

Table 11 shows that levels of TC were not significantly different between definite-FH and possible-FH group at Bonferroni test. This difference was statistically significant at Kruskal Wallis test (p=0.002) in all groups.

On the contrary levels of LDL-C gradually increased from patients without mutation to patients with a possible-FH and to patients with a definite-FH. Multiple comparisons by Bonferroni correction revealed statistically significant differences between each group versus the others (p=0.0026), Kruskal Wallis test revealed a significant difference between groups (p=0.001).

There was no significant difference in TG and HDL-C levels between the three groups

Table 11. Differences in lipid levels among groups based on the Simon Broome diagnosis

	no-FH	Significant at Bonferroni test	possible-FH	Significant at Bonferroni test	definie-FH	Significant at Kruskal Wallis
TC (mmol/L)	7.6 (6.6-8.7)	Sig vs pos p=0.030 Sig vs def p=0.002	8.1 (7.4-9.2)	Sig vs def p=0.088	8.5 (6.9-11.4)	p<0.0001
LDL–C (mmol/L)	5.4 (4.6-6.4)	Sig vs pos p=0.025 Sig vs def p<0.0001	6.3 (5.3-7.2)	Sig vs def p=0.026	6.8 (9.6-4.7)	p=0.001
HDL-C (mmol/L)	1.3 (1.2-1.7)	Sig vs pos p=ns Sig vs def p=0.014	1.3 (1.1-1.6)	Sig vs def p=0.080	1.1 (1.0-1.4)	p=0.039
no-HDL-C (mmol/L)	6.0 (4.9-6.8)	Sig vs pos p=0.014 Sig vs def p<0.0001	6.8 (5.8-7.7)	Sig vs def p=0.011	7.7 (5.5-10.0)	p<0.0001
TG (mmol/L)		Sig vs pos p=ns Sig vs def p=ns		Sig vs def p=ns		p=ns

4.7 Relation between diagnosis criteria and mutations

4.7.1 DCLN diagnosis and mutations

The number of patients in which mutations were found according to the DCLN score is show in figure 9. The highest number of positive patients was obtained when their clinical diagnosis was definite-FH. As expected, the frequency of FH-M^+ increased as increased the score from no FH group to possible, probable and to definite-FH. Patients without detected mutations were mainly in no FH group (59%) or in possible FH group (33%).
However, there was one patient with a definite-FH score but without any detected mutation.

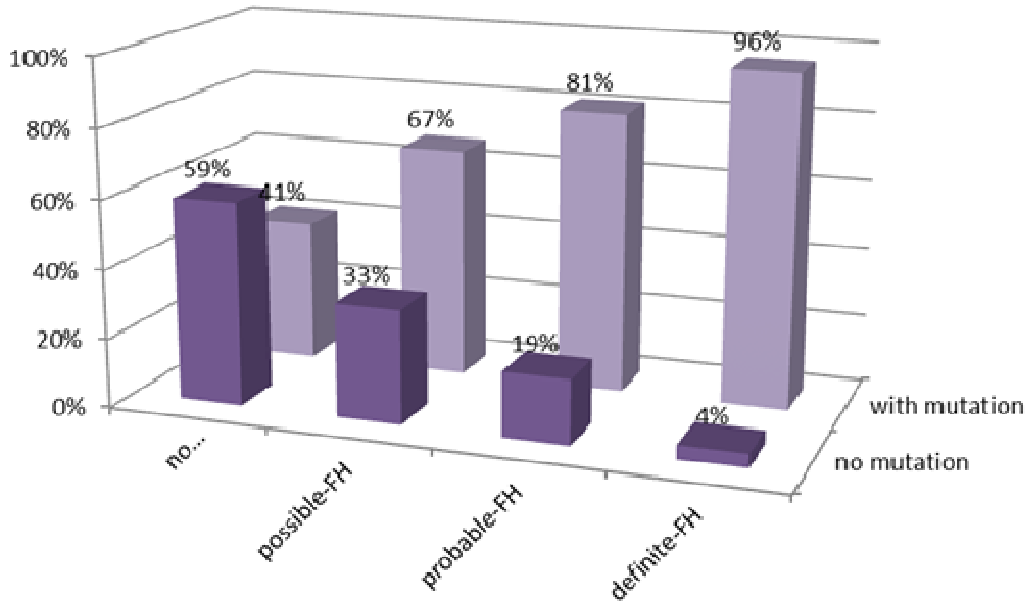

Figure. 9 Percentage of patients with or without mutation, classified by DCLN score

Furthermore, when mutations were considered on the basis of the mutation type (missense, radical homozygotes/compound heterozygote) we observed that the missense mutation had the same distribution among the four score groups (Fig.10).

In the groups of probable-FH and definite-FH we observed the highest frequency of the radical and homozygotes/compound heterozygote mutations that caused more severe phenotypes. However, patients (10%) with a radical mutation were present also in the no-FH group.
In the groups of possible-FH and no FH there was no the presence of patients with homozygotes/compound heterozygote mutations .

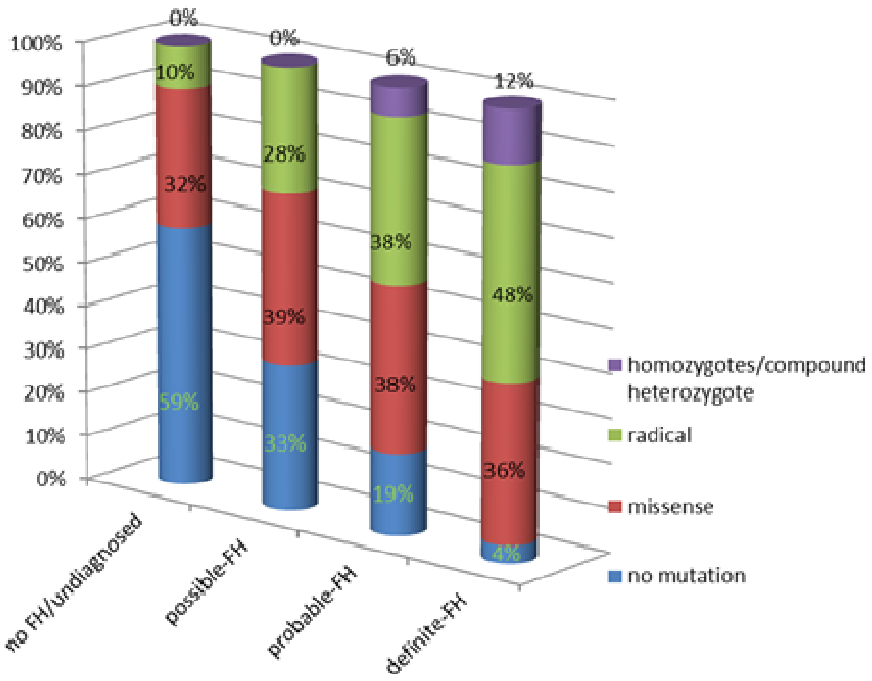

Figure 10. Distribution of mutation types according to DCLN score

4.7.2 Simon Broome diagnosis and mutations

The number of patients in which mutations were found according to the Simon Broome diagnostic criteria is show in figure 11. The highest number of positive patients was obtained when their clinical diagnosis was definite-FH. As expected, the frequency of FH-M^+ increased as increased the score from no FH group to possible, probable and to definite-FH. Patients without detected mutations were mainly in no FH group (49%) or in possible FH group (27%). However, in the group with uncertain diagnosis (no-FH) we found mutations in 51% .
This was possible because in this group we included patients with TC and LDL-C levels above the Simon Broome criteria cut off (7.5mmol/l and/or 4.9mmmol/l respectively) but without family history that could be clarified.

.

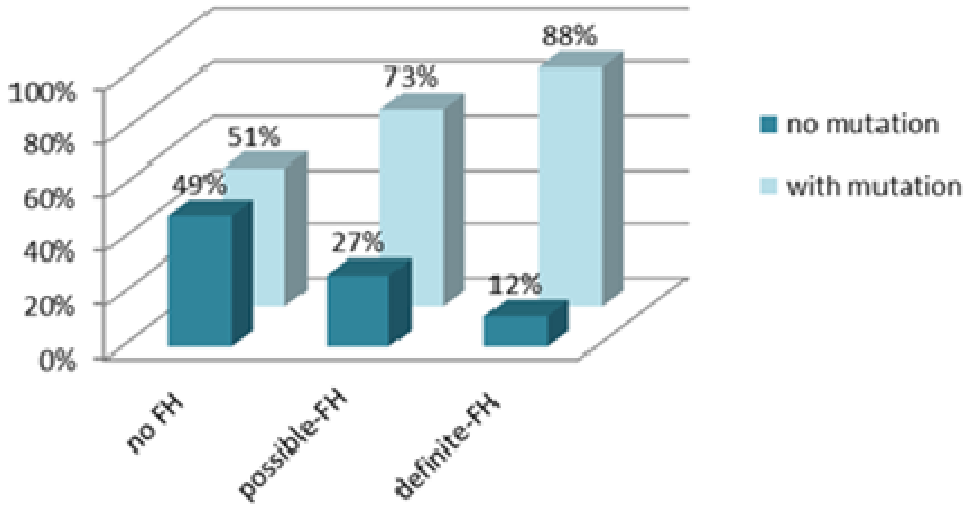

Figure11. Percentage of pa52tients with or without mutations, classified by Simon Broome criteria

Furthermore, when mutations were considered on the basis of mutation type (missense, radical, homozygotes/compound heterozygote) we observed that the missense mutations had the same distribution between possible-FH and definite –FH groups (Fig.12).
In the group of no-FH there were 27% of patients with missense mutation and 24% with radical mutation.
However there were no a homozygotes/compound heterozygote in the group of no-FH.

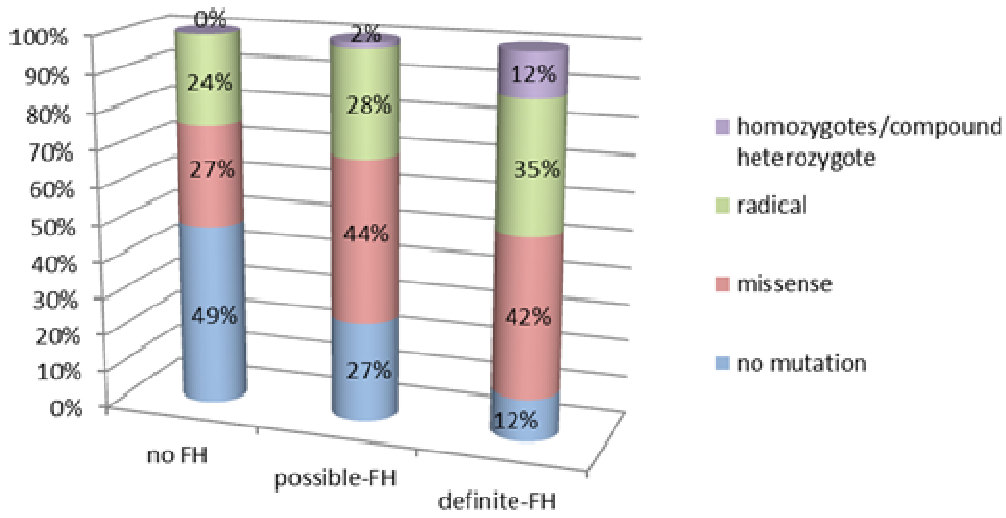

Figure 12. Distribution of mutation type according Simon Broome criteria

4.7.3 DCLN score vs Simon Broome criteria

The Chi-Quadrat Test analysis was used to compare the DCLN score and Simon Broome criteria. The two criteria were overlapping in the 68% of cases, showing a concordance of the 20.8% to classify the FH-M- in the no-FH group, a concordance of the 42.1% to classify the FH-M+ in the possible-FH group and of the 5.1% to classify the FH-M+ in the definite-FH group.
The differences were due to the presence of patients (16.1%) classified as probable/possible-FH on the base of the DCLN score but classified as no-FH on the base of the Simon Broome criteria.
Furthermore, there were two patients classified as definite-FH for DCLN score but no-FH for Simon Broome criteria. The patients show mutations in the LDLR gene. This difference in the classification could to be due to different clinical hallmarks of FH using the two diagnostic criteria. One patient presented corneal arcus and the other patient presented carotid atherosclerotic plaques. These clinical hallmarks of FH were not evaluated by Simon Broome criteria. For DCLN criteria the presence of corneal arcus assigns a score of 4 and the presence of carotid atherosclerotic plaques assigns a score of 2. These scores added to biochemical results allowed to establish the highest score of 8, classifying the patients with a definite-FH diagnosis.

Table 12 . Comparison between DCLN score and Simon Broome criteria

		Dutch Lipid Clinic Network		
Simon Broome criteria		no-FH	possible/probable-FH	definite-FH
	no-FH	53 (20.8%)	41 (16.1%)	2 (0.8%)
	possible-FH	20 (7.8%)	107 (42.1%)	10 (3.9%)
	definite-FH	0 (0)	8 (3.1%)	13 (5.1%)

4.8 *Polygenic*

4.8.1 LDL-C gene score analysis

We enrolled 199 patients with clinically diagnosed FH, of whom 160 were unrelated. The screening revealed mutations in 141 patients. The mean levels of TC and LDL-C was above 95th percentile (8.13 ± 2.12, 6.25 ± 2.13) respectively. Out of 199 FH samples (141mutation positive, 58 mutation negative), genotyped for all 12 SNPs, patients with polygenic FH with a gene score above 1.16 cut off were 10/141 (7.1%) FH/M^{+}, 5/58(8.6%) FH/M^{-}.

The familial hypercholesterolemia group without a known mutation had a significantly higher mean weighted LDL-C gene score of 0.96 (SD 0·21; p=0.0046) than did WHII participants. This suggests that a substantial proportion of the mutation-negative familial hypercholesterolemia group's raised LDL-C concentrations can be explained by co-inheritance of common LDL-C-raising SNPs.

When we calculated the weighted LDL-C gene score in the familial hypercholesterolemia group with a known mutation, the mean weighted score (0·93 [SD 0·20]) was significantly higher than the score in the WHII group (0·89 [0·23]; p=0.057, but was not significantly lower than that of the mutation-negative group (0.96 [0·21]) (Fig.13). This result suggests that even in patients with familial hypercholesterolemia who have a detected causative mutation, their raised LDL-C concentrations have an additional polygenic component.

Overall, 73 (16%) had a LDL-C gene score that fell within decile 10 of the WHII LDL-C gene score distribution, and 211 (46%) fell within deciles 7–10.

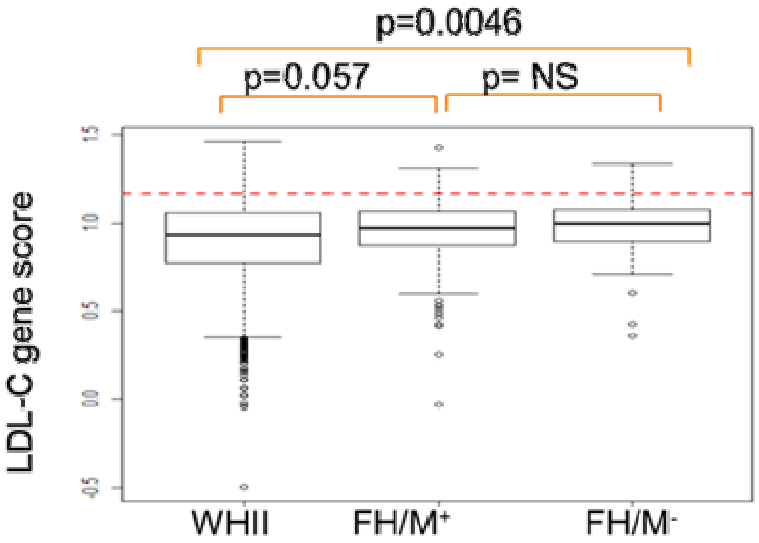

Figure 13. Comparison of the LDL-C SNPs score between the WHII control populations, the FH mutation positive individuals, the FH mutation negative individuals in a standard box-plot. Dashed red line indicates the top of decile cut off 1.16 for the WHII cohort.

4.9 *Paediatric cohort*

4.9.1 Patients characteristics

87 unrelated patients were included in the paediatric cohort (total populations 99 subjects). The mean age was 9.7 ± 3.8, the clinical characteristics of patients enrolled were reported in table 13.
All lipid parameters showed parametric distributions, the data was reported as mean±SD. The mean levels of total cholesterol (7.8mmol/l) and of LDL-cholesterol (5.7mmol/l) in FH-M^+ were significantly higher than FH-M^- ($p=0.002$, $p<0.001$ respectively). There was no significant difference in triglyceride levels between two groups. The clinical hallmarks showed no significant difference between the groups probably due to the presence of missing data.

Table 13 . Basic characteristics of paediatric FH patients

	Total n=99	**FH-M^+ n=65**	**FH-M^- n =22**	**Sig. M^+ vs M^-**
Age (years)	9.7 ± 3.8	9.5 ± 3.9	10.18± 3.5	ns
Total cholesterol (mmol/L)	7.8 ± 2.3	8.3 ± 2.4	6.4 ± 1.1	p=.0.002
LDL-cholesterol (mmol/L)	5.7 ± 2.1	6.3 ± 1.9	4.0 ± 1.6	p<0.0001
HDL-cholesterol (mmol/L)	1.3 ± 0.3	1.2 ± 0.2	1.5 ± 0.4	ns
no-HDL –Cholesterol (mmol/L)	6.3 ± 2.0	6.8 ± 2.0	4.9 ± 1.2	p<0.0001
Triglycerides (mmol/L)	0.9 (0.7-1.2)	0.9 (0.7-1.1)	1.0 (0.7-1.4)	ns
Corneal arcus	1 (2.5)	1 (3.3)	0 (0.0)	ns
Tendon xanthomas (%)	5 (8.5)	3 (6.5)	2 (15.4)	ns

Among 87 patients, there were 24 subjects no-FH (27.6%), 56 possible-FH (64.4%) and 7 with definite-FH (8%). in the group no-FH 42% of patients shows missense mutations and 21% show radical mutations but subjects with homozygotes/compound heterozygotes were absent. The missense mutation had a higher frequency in possible-FH than other groups. Among three patients with homozygotes/compound heterozygotes two were classified into possible-FH and one into definite-FH showing a frequency of 4% and 14% respectively (Table 14).
Table 15 showed that levels of TC and LDL-C gradually increase from patients without mutation to patients with a missense mutation, to patients

with a radical and to patients with homozygotes/compound heterozygotes. All these differences in TC and LDL-C were statistically significant at Kruskal Wallis test ($p<0.0001$).
Multiple testing by Bonferroni correction revealed statistical differences between each group versus the others. In particular the mean of TC levels in subjects with homozygotes/compound heterozygotes was double compared to patients with a missense mutation. There was no significant difference in TG and HCL-C levels between four groups.

Table 14. Frequency of mutation type among Simon Broome criteria in pediatric cohort

	No mutation	Missense	Radical	Homozygotes /compuond heterozygotes
no-FH	38%	42%	21%	0%
possible-FH	20%	41%	36%	4%
definite-FH	29%	29%	29%	14%

Table 15. Lipid levels differences among groups based on Simon Broome criteria in a pediatric cohort.

	no mutation n=108	Significant at Bonferroni	Missense n=117	Significant at Bonferroni	Radical n=87	Significant at Bonferroni	Homozygotes/ Compound heterozygotes n=10	Significant at Kruskal Wallis
TC (mmol/l)	6.4 ± 1.1	Sig vs Miss p=0.03 Sig vs Rad p<0.0001 Sig vs Comp p<0.0001	7.5 ± 1.2	ns vs Rad Sig vs Comp p<0.0001	8.3 ± 1.3	Sig vs Comp p<0.0001	16.6 ± 2.7	p<0.0001
LDL-C (mmol/l)	4.0 ± 1.6	Sig vs Miss p=0.001 Sig vs Rad p<0.0001 Sig vs Comp p<0.0001	5.7 ± 1.4	ns vs Rad Sig vs Comp p<0.0001	6.5 ± 1.2	Sig vs Comp p<0.0001	13.1 ± 0.8	p<0.0001
HDL-C (mmol/l)	1.5 ± 0.4	ns vs Miss ns vs Rad ns vs Comp	1.3 ± 0.2	ns vs Rad ns vs Comp	1.2 ±0.2	ns vs Comp	1.5 ± 0.4	p=0.03
No-HDL-C (mmmol)	1.5 ± 0.4	ns vs Miss Sig vs Rad p<0.0001 Sig vs Comp p<0.0001	1.3 ± 0.2	ns vs Rad Sig vs Comp p<0.0001	1.2 ± 0.2	Sig vs Comp p<0.0001	1.0 ±0.1	p<0.0001
TG (mmol/l)	1.0 (0.7-1.4)	ns vs Miss ns vs Rad ns vs Comp	0.8 (0.6-1.1)	ns vs Rad ns vs Comp	0.8 (0.7-1.1)	ns vs Comp	0.9 (0.9-0.9)	ns

5. Discussion

In the attempt to identify genetic background among patients with clinically diagnosed FH, 322 unrelated patients were screened for LDLR, APOB and PCSK9 genes, according to studies that underlined how the molecular analysis plays a key role to confirm the diagnosis. The identification of individuals with FH, as early as possible, is important to improve their therapeutic approach and drug monitoring, in fact subjects between 20 and 39 years old have 100-fold increased risk to develop premature CHD compared to the normal population.

The main features of genetic FH are the high mutation heterogeneity and presence of mutations clusters across different geographic areas.

In our cohort, the mutations have been found in 214/322 subjects with a mutation rate of 66% in agreement with other studies ([78], [79].

Out of 214 mutated subjects about 95.3% were carriers of *LDLR* mutations, 2.8% of *APOB* mutations and 1.8% of *PCSK9* mutations. The distribution of the mutations across the three candidate genes is similar to that reported in other Italian studies (97.4% *LDLR*, 2.2% in *APOB* and 0.36% in *PCSK9*) [82] while is little different to that reported in French cohort studies (73.9% *LDLR*, 6.6% in *APOB* and 0.7% in *PCSK9*) [83] and Dutch cohort studies (88% *LDLR*, 12% in *APOB* and no mutation in *PCSK9*) [84].

The main difference between these studies is the percentage and the type of mutation identified in *APOB* gene. The p.Arg3500G mutation is frequently detected in French and Dutch cohort. The Arg3500Gln mutation in the *APOB gene* is the most common cause of hypercholesterolemia, accounting for 5 to 10% of FH cases in Northern European populations [8] but its frequency is very low in the Latin European population compared with Central European population [35].

The genetic screening performed in FH patients demonstrated that *LDLR* gene shows a genetic heterogeneity in our population. Regarding the ApoB gene, through direct sequencing of exon 26 and 29 we identified four different mutations of which two novels (p.Val3306Ile and p.Trp3633Arg). The pathogenic role of the novel variations found in the APOB gene was supported by *in silico* analysis only for p.Val3306Ile. In a recent study Thomas et al [85] reported that a new mutation (p.Arg50Trp) identified in exon 3 of *APOB* causes hypercholesterolemia, underlining the importance to perform a complete screening for all the involved exons although the cost is still high.

In the PCSK9 gene we found four different variations (p.Pro331Ala, p.Arg499Hys novels) but we did not find the most frequent mutation (p.Asp374Tyr), which is associated with a higher CHD [86]. *In silico*

analysis of variants in PCSK9 suggested that its effect was uncertain. Variants could influence the functionality of the protein, although these algorithms cannot predict the GOF or LOF effect. In fact, it has been reported that predictions of GOF mutations should be interpreted with caution [87] due to the inaccuracy of the prediction tools. Functional characterization is necessary to evaluate the functional effect of these variants.

In *LDLR* the variations were detected along the entire gene. The UCL FH database (www.ucl.ac.uk) reported a different frequency of mutation on the exons. A greater number of different mutations were detected in exon 4 which is more susceptible to suffer modifications because of its length. However, the highest frequency of mutations in our cohort was identified in exon 10, exon 12, and exon 16. Mainly, the highest frequency of mutations in these exons is due to the presence of common *LDLR* gene mutations which are typical of southern Italian regions. The number of LDLR mutations accounting for the FH phenotype varies among European countries [88], [83]. In Italy, more than 237 mutations have been reported so far, some of which are present only in some Italian regions [82].

In Southern Italy the 6 most frequent mutations account for more than 50% of cases. The c.1775G>A mutation, most common in our cohort, has also been described in Greek and German FH patients [89]. The second most common mutation (13.7%) is c.2311+G>A, these data agree with previous reports, suggesting a founder effect in our region for this mutation [90].

Among the 58 different mutations, 15 have been found in our population for the first time, out of which 10 have been already published [91], [34] and 5 are novels.

The p.Cys34Trp mutation, is involved the ligand-binding domain. In particular, is one of six cysteine residues that forms LA repeat 4 that LDL-receptor used to bind APOE on chylomicrons. The p.Met298Val was found in compound heterozygotes together with splicing mutation p.Ala771_Ile796del.

The mutation p.Leu444Arg occurs at the same codon as that previously described for the c.1330T>C (p.Leu444Pro) mutation and it affects a highly conserved aminoacid located in the EGF precursor homologous domain [92].

Regarding mutation splicing c.694+1G>C, the analysis of RNA from the patient confirmed the less use of the donor splice site and the presence of a splicing alteration in intron 4. The transcript corresponding to the skipping of exon 4 was also identified in healthy subjects [93], [94] and in our control although at lower levels than the mutated sample. In addition, three bands not detectable in the healthy control were observed on the gel, which

suggests that the use of cryptic splice sites causes a dramatic decrease in the amount of full length mRNA.
The small duplication of AG, p.Cys358SerfsX13 is involved in the EGF –like domain. The EGF –like domain is responsible for the dissociation of LDLR from the LDL in the endosome at low pH, [26]. This mutation could to prevent recycling of the receptor to the cell membrane.
Although there is a high frequency of common mutations, we recommend performing the complete genetic screening aimed at identifying new or rare mutations as well as patients with two mutations. In fact, to prevent fatal cardiovascular events, particular attention should be paid to the identification of compound heterozygous and homozygous patients .
In addition, we examined the reliability of the Dutch Lipid Clinic Network (DCLN) score and Simon Broome criteria, to identify patients with a high or low probability of carrying an FH-causing mutation. Using the DCLN score, we found that the percentage of FH-M^+ classified as definite-FH is 96%, classified as probable-FH or possible-FH are 81.3% and 66.9% respectively and patients FH-M^+ classified as no-FH is 41.1%.
Using the Simon Broome criteria, we found that the percentage of FH-M^+ classified as definite-FH is 88.5%, classified as possible-FH are 73.2% and FH-M^+ classified as no-FH is 51.3%. Patients with mutations in the APOB gene are classified as no-FH for both diagnosis criteria. These results could be due to mutations in the APOB gene that show a less severe phenotype than the mutations in the LDLR gene [94].
The DCLN criteria are more accurate than Simon Broom showing very high sensitivity and specificity to detect patients with mutations.
Furthermore, the comparison of the two diagnostic criteria revealed an overlapping between possible-FH and possible-FH/probable-FH groups for both criteria. However we found two patients with mutations in the *LDLR* gene classified as definite-FH on the basis of DCLN score and no-FH on the basis of Simon Broome criteria. This result is probably due to the fact that the Simon Broome criteria do not take into account physical examinations such as corneal arcus, carotid atherosclerotic plaques and clinical history of CHD. Therefore, Simon Broome criteria are easy to apply and they can sometimes lead to a missed diagnosis.
Additionally, a proportion of patients (199 subjects) were analyzed for 12 common SNPs that the Global Lipid Genetic Consortium (GLGC) reported as significantly associated with LDL-C. The cumulative effect of common LDL-rising alleles in these genes was shown to be likely cause of high LDL-C in a proportion of examined patients. These results confirm the previous report by Talmud et al, and show that, also in Italian patients with the FH phenotype but without mutations in the main candidate genes, there is a

likely polygenic cause, due to the inheritance of LDL-C-raising SNPs which increases LDL-C concentration in patients. This polygenic contribution is also seen in patients with detected FH-causing mutations.
We examined also a cohort of children exposed to fewer environmental factors in terms of their lipoprotein metabolism in which an unequivocal diagnosis of FH is easier to establish. In this population we observed the gradually increase of the TC and LDL-C among patients with different types of mutation showing that the type of the LDLR mutation influences the lipid profile. Patients with radical mutations show a worse lipid profile than missense carriers allowing a prognostic evaluation for physicians.
In conclusion, the results of this study represent an update of FH genetic background in an Italian population of patients from southern regions.
These data enlarge the spectrum of mutations causing FH and highlight the importance of a complete screening of causative genes in order to identify more or novel mutations and to identify homozygotes/compound heterozygotes.
The correlation between mutation types and lipid profile underlines importance of genetic screening as a prognostic tool.
The presence of a percentage of mutated subjects in the no-FH group according to Dutch Lipid Clinic Network (DCLN) score and Simon Broome criteria suggests that genetic screening is also useful to confirm the diagnosis, especially in patients with an uncertain phenotype.

References

1. Soutar, A.K. and R.P. Naoumova, *Mechanisms of disease: genetic causes of familial hypercholesterolemia.* Nat Clin Pract Cardiovasc Med, 2007. **4**(4): p. 214-25.
2. Innerarity, T.L., et al., *Familial defective apolipoprotein B-100: a mutation of apolipoprotein B that causes hypercholesterolemia.* J Lipid Res, 1990. **31**(8): p. 1337-49.
3. Abifadel, M., et al., *Mutations in PCSK9 cause autosomal dominant hypercholesterolemia.* Nat Genet, 2003. **34**(2): p. 154-6.
4. Varret, M., et al., *Genetic heterogeneity of autosomal dominant hypercholesterolemia.* Clin Genet, 2008. **73**(1): p. 1-13.
5. Wang, X., et al., *Genome-wide linkage scan of a pedigree with familial hypercholesterolemia suggests susceptibility loci on chromosomes 3q25-26 and 21q22.* PLoS One, 2011. **6**(10): p. e24838.
6. Soutar, A.K. and R.P. Naoumova, *Autosomal recessive hypercholesterolemia.* Semin Vasc Med, 2004. **4**(3): p. 241-8.
7. Hobbs, H.H., M.S. Brown, and J.L. Goldstein, *Molecular genetics of the LDL receptor gene in familial hypercholesterolemia.* Hum Mutat, 1992. **1**(6): p. 445-66.
8. Hopkins, P.N., et al., *Familial hypercholesterolemias: prevalence, genetics, diagnosis and screening recommendations from the National Lipid Association Expert Panel on Familial Hypercholesterolemia.* J Clin Lipidol, 2011. **5**(3 Suppl): p. S9-17.
9. Farnier, M. and E. Bruckert, *Severe familial hypercholesterolaemia: current and future management.* Arch Cardiovasc Dis, 2012. **105**(12): p. 656-65.
10. Raal, F.J., et al., *Reduction in mortality in subjects with homozygous familial hypercholesterolemia associated with advances in lipid-lowering therapy.* Circulation, 2011. **124**(20): p. 2202-7.
11. Hovingh, G.K., et al., *Diagnosis and treatment of familial hypercholesterolaemia.* Eur Heart J, 2013. **34**(13): p. 962-71.
12. van der Graaf, A., et al., *Molecular basis of autosomal dominant hypercholesterolemia: assessment in a large cohort of hypercholesterolemic children.* Circulation, 2011. **123**(11): p. 1167-73.
13. Marks, D., et al., *A review on the diagnosis, natural history, and treatment of familial hypercholesterolaemia.* Atherosclerosis, 2003. **168**(1): p. 1-14.
14. Austin, M.A., et al., *Genetic causes of monogenic heterozygous familial hypercholesterolemia: a HuGE prevalence review.* Am J Epidemiol, 2004. **160**(5): p. 407-20.
15. Sullivan, D., *Guidelines for the diagnosis and management of familial hypercholesterolaemia.* Heart Lung Circ, 2007. **16**(1): p. 25-7.
16. Civeira, F., *Guidelines for the diagnosis and management of heterozygous familial hypercholesterolemia.* Atherosclerosis, 2004. **173**(1): p. 55-68.

17. Minhas, R., et al., *Controversies in familial hypercholesterolaemia: recommendations of the NICE Guideline Development Group for the identification and management of familial hypercholesterolaemia.* Heart, 2009. **95**(7): p. 584-7; discussion 587-91.
18. Niu, D.M., et al., *Clinical observations, molecular genetic analysis, and treatment of sitosterolemia in infants and children.* J Inherit Metab Dis, 2010. **33**(4): p. 437-43.
19. Nordestgaard, B.G., et al., *Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society.* Eur Heart J, 2013. **34**(45): p. 3478-90a.
20. van Aalst-Cohen, E.S., et al., *Diagnosing familial hypercholesterolaemia: the relevance of genetic testing.* Eur Heart J, 2006. **27**(18): p. 2240-6.
21. Kusters, D.M., et al., *Design and baseline data of a pediatric study with rosuvastatin in familial hypercholesterolemia.* J Clin Lipidol, 2013. **7**(5): p. 408-13.
22. Daniels, S.R., S.S. Gidding, and S.D. de Ferranti, *Pediatric aspects of familial hypercholesterolemias: recommendations from the National Lipid Association Expert Panel on Familial Hypercholesterolemia.* J Clin Lipidol, 2011. **5**(3 Suppl): p. S30-7.
23. Langslet, G. and L. Ose, *Screening methods in the diagnosis and assessment of children and adolescents with familial hypercholesterolemia.* Expert Rev Cardiovasc Ther, 2013. **11**(8): p. 1061-6.
24. Bates, T.R., et al., *Detection of familial hypercholesterolaemia: a major treatment gap in preventative cardiology.* Heart Lung Circ, 2008. **17**(5): p. 411-3.
25. Goldstein, J.L. and M.S. Brown, *Binding and degradation of low density lipoproteins by cultured human fibroblasts. Comparison of cells from a normal subject and from a patient with homozygous familial hypercholesterolemia.* J Biol Chem, 1974. **249**(16): p. 5153-62.
26. Brown, M.S. and J.L. Goldstein, *A receptor-mediated pathway for cholesterol homeostasis.* Science, 1986. **232**(4746): p. 34-47.
27. Goldstein, J.L., et al., *Receptor-mediated endocytosis: concepts emerging from the LDL receptor system.* Annu Rev Cell Biol, 1985. **1**: p. 1-39.
28. Ranheim, T., et al., *Model system for phenotypic characterization of sequence variations in the LDL receptor gene.* Clin Chem, 2006. **52**(8): p. 1469-79.
29. Usifo, E., et al., *Low-density lipoprotein receptor gene familial hypercholesterolemia variant database: update and pathological assessment.* Ann Hum Genet, 2012. **76**(5): p. 387-401.
30. Schmitz, G., et al., *Fluorescence flow cytometry of human leukocytes in the detection of LDL receptor defects in the differential diagnosis of hypercholesterolemia.* Arterioscler Thromb Vasc Biol, 1993. **13**(7): p. 1053-1065.

31. Chan, P., et al., *Surface expression of low density lipoprotein receptor in EBV-transformed lymphocytes: characterization and use for studying familial hypercholesterolemia.* Atherosclerosis, 1997. **131**(2): p. 149-60.
32. Raungaard, B., et al., *Flow cytometric assessment of LDL receptor activity in peripheral blood mononuclear cells compared to gene mutation detection in diagnosis of heterozygous familial hypercholesterolemia.* Cytometry, 1999. **36**(1): p. 52-9.
33. Urdal, P., et al., *Flow cytometric measurement of low density lipoprotein receptor activity validated by DNA analysis in diagnosing heterozygous familial hypercholesterolemia.* Cytometry, 1997. **30**(5): p. 264-8.
34. Romano, M., et al., *An improved method on stimulated T-lymphocytes to functionally characterize novel and known LDLR mutations.* J Lipid Res, 2011. **52**(11): p. 2095-100.
35. Liyanage, K.E., et al., *Familial hypercholesterolemia: epidemiology, Neolithic origins and modern geographic distribution.* Crit Rev Clin Lab Sci, 2011. **48**(1): p. 1-18.
36. Bertolini, S., et al., *Clinical Expression of Familial Hypercholesterolemia in Clusters of Mutations of the LDL Receptor Gene That Cause a Receptor-Defective or Receptor-Negative Phenotype.* Arterioscler Thromb Vasc Biol, 2000. **20**(9): p. e41-e52.
37. Futema, M., et al., *Use of targeted exome sequencing as a diagnostic tool for Familial Hypercholesterolaemia.* J Med Genet, 2012. **49**(10): p. 644-9.
38. Kane, J.P., *Apolipoprotein B: structural and metabolic heterogeneity.* Annu Rev Physiol, 1983. **45**: p. 637-50.
39. Contois, J.H., et al., *Apolipoprotein B and cardiovascular disease risk: position statement from the AACC Lipoproteins and Vascular Diseases Division Working Group on Best Practices.* Clin Chem, 2009. **55**(3): p. 407-19.
40. Marz, W., et al., *Accumulation of "small dense" low density lipoproteins (LDL) in a homozygous patients with familial defective apolipoprotein B-100 results from heterogenous interaction of LDL subfractions with the LDL receptor.* J Clin Invest, 1993. **92**(6): p. 2922-33.
41. Boren, J., et al., *The molecular mechanism for the genetic disorder familial defective apolipoprotein B100.* J Biol Chem, 2001. **276**(12): p. 9214-8.
42. Goldstein, J.L., et al., *Hyperlipidemia in coronary heart disease. II. Genetic analysis of lipid levels in 176 families and delineation of a new inherited disorder, combined hyperlipidemia.* J Clin Invest, 1973. **52**(7): p. 1544-68.
43. Vrablik, M., R. Ceska, and A. Horinek, *Major apolipoprotein B-100 mutations in lipoprotein metabolism and atherosclerosis.* Physiol Res, 2001. **50**(4): p. 337-43.
44. Pullinger, C.R., et al., *Familial ligand-defective apolipoprotein B. Identification of a new mutation that decreases LDL receptor binding affinity.* J Clin Invest, 1995. **95**(3): p. 1225-34.

45. Tarugi, P. and M. Averna, *Hypobetalipoproteinemia: genetics, biochemistry, and clinical spectrum.* Adv Clin Chem, 2011. **54**: p. 81-107.
46. Burnett, J.R., et al., *A novel nontruncating APOB gene mutation, R463W, causes familial hypobetalipoproteinemia.* J Biol Chem, 2003. **278**(15): p. 13442-52.
47. Walldius, G., et al., *High apolipoprotein B, low apolipoprotein A-I, and improvement in the prediction of fatal myocardial infarction (AMORIS study): a prospective study.* The Lancet, 2001. **358**(9298): p. 2026-2033.
48. Sankatsing, R.R., et al., *Hepatic and cardiovascular consequences of familial hypobetalipoproteinemia.* Arterioscler Thromb Vasc Biol, 2005. **25**(9): p. 1979-84.
49. Benjannet, S., et al., *The proprotein convertase (PC) PCSK9 is inactivated by furin and/or PC5/6A: functional consequences of natural mutations and post-translational modifications.* J Biol Chem, 2006. **281**(41): p. 30561-72.
50. Lopez, D., *PCSK9: an enigmatic protease.* Biochim Biophys Acta, 2008. **1781**(4): p. 184-91.
51. Piper, D.E., et al., *The crystal structure of PCSK9: a regulator of plasma LDL-cholesterol.* Structure, 2007. **15**(5): p. 545-52.
52. Zhang, D.W., et al., *Structural requirements for PCSK9-mediated degradation of the low-density lipoprotein receptor.* Proc Natl Acad Sci U S A, 2008. **105**(35): p. 13045-50.
53. Zhang, D.W., et al., *Binding of proprotein convertase subtilisin/kexin type 9 to epidermal growth factor-like repeat A of low density lipoprotein receptor decreases receptor recycling and increases degradation.* J Biol Chem, 2007. **282**(25): p. 18602-12.
54. Shan, L., et al., *PCSK9 binds to multiple receptors and can be functionally inhibited by an EGF-A peptide.* Biochem Biophys Res Commun, 2008. **375**(1): p. 69-73.
55. Poirier, S., et al., *Dissection of the endogenous cellular pathways of PCSK9-induced low density lipoprotein receptor degradation: evidence for an intracellular route.* J Biol Chem, 2009. **284**(42): p. 28856-64.
56. Cameron, J., et al., *Effect of mutations in the PCSK9 gene on the cell surface LDL receptors.* Hum Mol Genet, 2006. **15**(9): p. 1551-8.
57. Cohen, J.C., et al., *Sequence variations in PCSK9, low LDL, and protection against coronary heart disease.* N Engl J Med, 2006. **354**(12): p. 1264-72.
58. He, G., et al., *ARH is a modular adaptor protein that interacts with the LDL receptor, clathrin, and AP-2.* J Biol Chem, 2002. **277**(46): p. 44044-9.
59. Chen, W.J., J.L. Goldstein, and M.S. Brown, *NPXY, a sequence often found in cytoplasmic tails, is required for coated pit-mediated internalization of the low density lipoprotein receptor.* J Biol Chem, 1990. **265**(6): p. 3116-23.
60. Obermeier, A., et al., *Neuronal differentiation signals are controlled by nerve growth factor receptor/Trk binding sites for SHC and PLC gamma.* EMBO J, 1994. **13**(7): p. 1585-90.

61. Fellin, R., et al., *Clinical and biochemical characterisation of patients with autosomal recessive hypercholesterolemia (ARH).* Nutr Metab Cardiovasc Dis, 2003. **13**(5): p. 278-86.
62. Pisciotta, L., et al., *Autosomal recessive hypercholesterolemia (ARH) and homozygous familial hypercholesterolemia (FH): a phenotypic comparison.* Atherosclerosis, 2006. **188**(2): p. 398-405.
63. Arca, M., et al., *Autosomal recessive hypercholesterolaemia in Sardinia, Italy, and mutations in ARH: a clinical and molecular genetic analysis.* Lancet, 2002. **359**(9309): p. 841-7.
64. Rahalkar, A.R. and R.A. Hegele, *Monogenic pediatric dyslipidemias: classification, genetics and clinical spectrum.* Mol Genet Metab, 2008. **93**(3): p. 282-94.
65. Kathiresan, S., et al., *Common variants at 30 loci contribute to polygenic dyslipidemia.* Nat Genet, 2009. **41**(1): p. 56-65.
66. Teslovich, T.M., et al., *Biological, clinical and population relevance of 95 loci for blood lipids.* Nature, 2010. **466**(7307): p. 707-13.
67. Talmud, P.J., et al., *Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolaemia: a case-control study.* Lancet, 2013. **381**(9874): p. 1293-301.
68. Broekhuizen, K., et al., *Is the process of delivery of an individually tailored lifestyle intervention associated with improvements in LDL cholesterol and multiple lifestyle behaviours in people with familial hypercholesterolemia?* BMC Public Health, 2012. **12**: p. 348.
69. *Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group.* N Engl J Med, 1998. **339**(19): p. 1349-57.
70. *Mortality in treated heterozygous familial hypercholesterolaemia: implications for clinical management. Scientific Steering Committee on behalf of the Simon Broome Register Group.* Atherosclerosis, 1999. **142**(1): p. 105-12.
71. Rozman, D. and K. Monostory, *Perspectives of the non-statin hypolipidemic agents.* Pharmacol Ther, 2010. **127**(1): p. 19-40.
72. Thompson, G.R., et al., *Severe hypercholesterolaemia: therapeutic goals and eligibility criteria for LDL apheresis in Europe.* Curr Opin Lipidol, 2010. **21**(6): p. 492-8.
73. Espinheira Mdo, C., et al., *Hypercholesterolemia--a disease with expression from childhood.* Rev Port Cardiol, 2013. **32**(5): p. 379-86.
74. Goldberg, A.C., *Novel therapies and new targets of treatment for familial hypercholesterolemia.* J Clin Lipidol, 2010. **4**(5): p. 350-6.
75. Watts, G.F., et al., *Integrated guidance on the care of familial hypercholesterolaemia from the International FH Foundation.* Int J Cardiol, 2014. **171**(3): p. 309-25.

76. Lilly, S.M. and D.J. Rader, *New targets and emerging therapies for reducing LDL cholesterol.* Curr Opin Lipidol, 2007. **18**(6): p. 650-5.
77. Sjouke, B., et al., *Familial hypercholesterolemia: present and future management.* Curr Cardiol Rep, 2011. **13**(6): p. 527-36.
78. Farnier, M., *PCSK9: From discovery to therapeutic applications.* Arch Cardiovasc Dis, 2014. **107**(1): p. 58-66.
79. Wierzbicki, A.S., et al., *New therapies to reduce low-density lipoprotein cholesterol.* Curr Opin Cardiol, 2013. **28**(4): p. 452-7.
80. Abravaya, K., et al., *Molecular beacons as diagnostic tools: technology and applications.* Clin Chem Lab Med, 2003. **41**(4): p. 468-74.
81. Ruotolo, A., et al., *The novel variant p.Ser465Leu in the PCSK9 gene does not account for the decreased LDLR activity in members of a FH family.* Clin Chem Lab Med, 2014.
82. Bertolini, S., et al., *Spectrum of mutations and phenotypic expression in patients with autosomal dominant hypercholesterolemia identified in Italy.* Atherosclerosis, 2013. **227**(2): p. 342-8.
83. Marduel, M., et al., *Molecular spectrum of autosomal dominant hypercholesterolemia in France.* Hum Mutat, 2010. **31**(11): p. E1811-24.
84. Kusters, D.M., et al., *Founder mutations in the Netherlands: geographical distribution of the most prevalent mutations in the low-density lipoprotein receptor and apolipoprotein B genes.* Neth Heart J, 2011. **19**(4): p. 175-182.
85. Thomas, E.R., et al., *Identification and biochemical analysis of a novel APOB mutation that causes autosomal dominant hypercholesterolemia.* Mol Genet Genomic Med, 2013. **1**(3): p. 155-61.
86. Pecin, I., et al., *Mutation detection in Croatian patients with familial hypercholesterolemia.* Ann Hum Genet, 2013. **77**(1): p. 22-30.
87. Flanagan, S.E., A.M. Patch, and S. Ellard, *Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations.* Genet Test Mol Biomarkers, 2010. **14**(4): p. 533-7.
88. Medeiros, A.M., et al., *Update of the Portuguese Familial Hypercholesterolaemia Study.* Atherosclerosis, 2010. **212**(2): p. 553-8.
89. Diakou, M., et al., *Spectrum of LDLR gene mutations, including a novel mutation causing familial hypercholesterolaemia, in North-western Greece.* Eur J Intern Med, 2011. **22**(5): p. e55-9.
90. Liguori, R., et al., *LDL receptor cDNA sequence analysis in familial hypercholesterolemia patients: 5 novel mutations with high prevalence in families originating from southern Italy.* Hum Mutat, 2001. **17**(5): p. 433.
91. Romano, M., et al., *Identification and functional characterization of LDLR mutations in familial hypercholesterolemia patients from Southern Italy.* Atherosclerosis, 2010. **210**(2): p. 493-6.

92. Fouchier, S.W., J.J. Kastelein, and J.C. Defesche, *Update of the molecular basis of familial hypercholesterolemia in The Netherlands.* Hum Mutat, 2005. **26**(6): p. 550-6.
93. Holla, O.L., et al., *Effects of intronic mutations in the LDLR gene on pre-mRNA splicing: Comparison of wet-lab and bioinformatics analyses.* Mol Genet Metab, 2009. **96**(4): p. 245-52.
94. Tveten, K., et al., *Analysis of alternatively spliced isoforms of human LDL receptor mRNA.* Clin Chim Acta, 2006. **373**(1-2): p. 151-7.

Work experience

07-05-2013 05-08-2013 I attend the laboratory of Prof Steve Humphries. Insitute of Cardiovascular Science, Centre for Cardiovascular Genetics at University College London.

Original papers

Romano M, Di Taranto MD, **D'Agostino MN**, Marotta G, Gentile M, Abate G, Mirabelli P, Di Noto R, Del Vecchio L, Rubba P, Fortunato G. Identification and functional characterization of LDLR mutations in familial hypercholesterolemia patients from Southern Italy. Atherosclerosis. 2010 Jun;210(2):493-6.

Romano M, Di Taranto MD, Mirabelli P, **D'Agostino MN**, Iannuzzi A, Marotta G, Gentile M, Raia M, Di Noto R, Del Vecchio L, Rubba P, Fortunato G. An improved method on stimulated T-lymphocytes to functionally characterize novel and known LDLR mutations. J Lipid Res. 2011 Nov;52(11):2095-100.

Antonino Staiano, Maria Donata Di Taranto, Elena Bloise, **Maria Nicoletta D'Agostino**, Antonietta D'Angelo, Gennaro Marotta, Marco Gentile, Fabrizio Jossa, Arcangelo Iannuzzi, Paolo Rubba, Giuliana Fortunato. Investigation of Single Nucleotide Polymorphisms Associated to Familial Combined Hyperlipidemia with Random Forests. Neural Nets and Surroundings. Smart Innovation, Systems and Technologies Book Series. 2013;19:169-178.

Ruotolo A, Di Taranto MD, **D'Agostino MN**, Marotta G, Gentile M, Nunziata M, Sodano M, Di Noto R, Del Vecchio L, Rubba P, Fortunato G. The novel variant p.Ser465Leu in the PCSK9 gene does not account for the decreased LDLR activity in members of a FH family. Clin Chem Lab Med. 2014 Mar 6. [Epub ahead of print].

Speaker at congress

Alterazioni nel recettore LDL nell'Iperlipidemia Combinata Familiare. Congresso Regionale Società Italiana per lo Studio dell'Arteriosclerosi Sessione Campania Naploli (Italy)., September 5, 2012.

Study of polygenic familial hypercholesterolemia in Southern Italy 27° Congresso Nazionale Società Italiana per lo Studio dell'Arteriosclerosi. Roma (Italy); november 27-29 2013. Abstract winner of the "Travel Grant"

Abstracts at congresses

M.D. Di Taranto, **M.N. D'Agostino**, E. Bloise, M. Romano, G. Marotta, M. Gentile, A. Iannuzzi, P. Rubba, G. Fortunato. Familial Combined Hyperlipidemia: identification of misdiagnosed patients by detection of LDLR mutations. 36th FEBS Congress. Torino (Italy); June 25-30, 2011.

M.N. D'Agostino, M. Romano, M.D. Di Taranto, G. Marotta, M. Gentile, A. Iannuzzi, P. Rubba, G. Fortunato. Molecular diagnosis of Familial Hypercholesterolemia: the utility of a country specific protocol. 79th European Atherosclerosis Society Congress. Gotheburg, Svezia; June 26-29, 2011.

M.D. Di Taranto, E. Bloise, **M.N. D'Agostino**, A. D'Angelo,G. Marotta, M. Gentile, F. Jossa, A. Iannuzzi, P. Rubba, G.Fortunato, A. Staiano. SNP associated to Familial Combined Hyperlipidemia: application of an ensemble of decision trees protocol. 25° Congresso Nazionale Società Italiana per lo Studio della Arteriosclerosi. Roma; 30 November to 3 December, 2011. Abstract winner of the "Travel Grant"

M.N. D'Agostino, M.D. Di Taranto, E. Bloise, A. D'Angelo, I. Altieri, G. Marotta, M. Gentile, F. Jossa, A. Iannuzzi, P.Rubba, G. Fortunato. Identification of LDLR mutations in patients with Familial Combined Hyperlipidemia. 25° Congresso Nazionale Società Italiana per lo Studio della Arteriosclerosi. Roma; 30 November to 3 December, 2011. Abstract winner of the "Travel Grant"

A. Ruotolo, M. Romano, M.D. Di Taranto, P. Mirabelli, **M.N. D'Agostino**, A. Iannuzzi, G. Marotta, M. Gentile, R. Di Noto, L. Del Vecchio, P. Rubba, G. Fortunato. Identification and functional characterization of a new mutation leading to defective uptake of LDL-LDLR complex. 25° Congresso Nazionale Società Italiana per lo Studio della Arteriosclerosi.
Roma; 30 November to 3 December, 2011. Abstract winner of the "Travel Grant"

M. Nunziata, G. Marotta, M.G. Coppola, M. F. Coppola, M.Gentile, G. Fortunato, **M.N. D'Agostino**, M.D. Di Taranto, P. Rubba. Management of severe hypercholesterolemia in a family bearing c.974G>A LDL receptor mutation. 25° Congresso Nazionale Società Italiana per lo Studio della Arteriosclerosi. Roma; 30 November to 3 December, 2011. Abstract winner of the "Travel Grant"

M.G. Coppola, F. Jossa, M. Nunziata, M. Gentile, G. De Simone, M. De Marco, G. Fortunato, **M.N. D'Agostino**, M.D. Di Taranto, G. Marotta, P. Rubba. Echocardiografic and ecodoppler abnormalities in relation to LDL cholesterol in Familial Hypercholesterolemia. 25° Congresso Nazionale Società Italiana per lo Studio della Arteriosclerosi. Roma; 30 November to 3 December, 2011. Abstract winner of the "Travel Grant"

M. Sodano, G.Marotta, **M.N. D'Agostino**, M. Romano, M. Gentile, G. Fortunato, P. Rubba. LDL receptor alterations in a patient with clinical diagnosis of familial combined hyperlipidemia. 25° Congresso Nazionale Società Italiana per lo Studio della Arteriosclerosi. Roma; 30 November - 3 Dicember 2011. Giornale Italiano dell'Aterosclerosi 2011; 2(3): 97

A. Staiano, M.D. Di Taranto, E. Bloise, **M.N. D'Agostino**, A. D'Angelo, A. Morgante, G. Marotta, M. Gentile, F. Jossa, A. Iannuzzi, P. Rubba, G. Fortunato. Genetic of Familial Combined Hyperlipidemia: different approaches to study association of Single Nucleotide Polymorphisms. 80th European Atherosclerosis Society Congress. Milano; May 25-28, 2012.

A. D'Angelo, **M. N. D'Agostino**, M. D. Di Taranto, A. Ruotolo, G. Marotta, M. Gentile, P. Rubba,and G. Fortunato. Alterations Of The LDL Receptor Gene In Familial Combined Hyperlipidemia. 44° Congresso Nazionale Società Italiana di Biochimica Clinica e Biologia Molecolare Clinica. Roma (Italy); November 5-7 2012.

A. Ruotolo, **M. N. D'Agostino**, M. D. Di Taranto, A. D'Angelo, G. Marotta, M. Gentile, P. Rubba, and G. Fortunato, Familial Hypercholesterolemia: identification of a new mutation in the PCSK9 gene. 44° Congresso Nazionale Societa' Italiana di Biochimica Clinica e Biologia Molecolare Clinica. Roma (Italy); November 5-7 2012

M. N. D'Agostino, M. D. Di Taranto, A. D'Angelo, A.Ruotolo, G. Marotta, M. Gentile , F. Jossa , A. Iannuzzi, P. Rubba, G. Fortunato. Familial Hypercholesterolemia: An update of the genetic characterization. 26° Congresso Nazionale Società Italiana per lo Studio dell'Arteriosclerosi. Roma (Italy); November 28-30 2012. Abstract winner of the "Travel Grant"

A. Ruotolo, **M. N. D'Agostino**, M. D.Di Taranto , A. D'Angelo, G. Marotta , M. Gentile, F. Jossa, P. Rubba, G. Fortunato. Identification of a new mutation in the PCSK9 gene in a patient with Familial Hypercholesterolemia. 26° Congresso Nazionale Società Italiana per lo Studio dell'Arteriosclerosi. Roma (Italy); November 28-30 2012. Abstract winner of the "Travel Grant"

M.G. Coppola, M. Nunziata, G. Marotta, F. Jossa, M. Gentile, G. Fortunato**, M.N. D'Agostino**, M.D. Di Taranto, P. Rubba. Aortic stenosis in heterozygous familial hypercholesterolemia with IVS 10+1 G>A LDL receptor mutation: a case report. 26° Congresso Nazionale Società Italiana per lo Studio dell'Arteriosclerosi. Roma (Italy); November 28-30 2012. Abstract winner of the "Travel Grant"

D'Angelo A, **D'Agostino MN**, Di Taranto MD, Ruotolo A, Giacobbe C, Marotta G, Gentile M, Jossa F, Iannuzzi A, Rubba P, Fortunato G. Genetic screening of Familial Hypercholesterolemia in Southern Italy: a proposed diagnostic flowchart. 20th IFCC-EFLM European Congress of Clinical Chemistry and Laboratory Medicine (EuroMedLab) e 45th Congress of the Italian Society of Clinical Biochemistry and Clinical Molecular Biology (SIBioC). Milano; 19-23 May 2013.

M.N. D'Agostino, M. D. Di Taranto, A. D'Angelo, A. Ruotolo, C. Giacobbe, G. Marotta, M. Gentile, P. Rubba, and G. Fortunato. Familial Combined Hyperlipidemia: identification of LDLR mutations in a population from Southern Italy. 81th European Atherosclerosis Society Congress. Lyon (France); June 2-5, 2013.

Giacobbe G, Melis D, Di Taranto MD, **D'Agostino MN**, Di Fiore R,Parenti G, Minopoli G, Andria G, Fortunato G. ApoE genotypes are correlated with levels of triglyceride and cholesterol in patients with glycogen storage diseases types 1a and 1b. XVI Congresso Nazionale della Società Italiana di Genetica Umana. Roma; 25-28 September2013

Ruotolo A, **D'Agostino MN**, Di Taranto MD, D'Angelo A, Pagano C, Marotta, Gentile M, Giuliana Fortunato. Screening of APOB and PCSK9 genes in Familial Hypercholesterolemia patients without mutations in the LDLR gene. 27° Congresso Nazionale Società Italiana per lo Studio della Arteriosclerosi. Roma; 27-29 November 2013. Abstract winner of the "Travel Grant"

D'Agostino MN, Futema M, Li K, D'Angelo A, Rubba P, Fortunato G, Humphries SE. Study of polygenic Familial Hypercholesterolemia in Southern Italy. 7° Congresso Nazionale Società Italiana per lo Studio della Arteriosclerosi. Roma; 7-9 November 2013. Abstract winner of the "Travel Grant"

Atherosclerosis 210 (2010) 493–496

Contents lists available at ScienceDirect

Atherosclerosis

journal homepage: www.elsevier.com/locate/atherosclerosis

Short communication

Identification and functional characterization of LDLR mutations in familial hypercholesterolemia patients from Southern Italy

Maria Romano[a,b], Maria Donata Di Taranto[c], Maria Nicoletta D'Agostino[a,b], Gennaro Marotta[d], Marco Gentile[d], Giovanna Abate[a], Peppino Mirabelli[a,b], Rosa Di Noto[a,b], Luigi Del Vecchio[a,b], Paolo Rubba[d], Giuliana Fortunato[a,b,*]

[a] CEINGE S.C.a r.l. Biotecnologie Avanzate, Napoli, Italy
[b] Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, Napoli, Italy
[c] Fondazione IRCCS SDN, Napoli, Italy
[d] Dipartimento di Medicina Clinica e Sperimentale, Università degli Studi di Napoli Federico II, Napoli, Italy

ARTICLE INFO

Article history:
Received 20 November 2009
Accepted 29 November 2009
Available online 5 December 2009

Keywords:
LDL receptor
Familial hypercholesterolemia
Large rearrangements
LDL uptake assay

ABSTRACT

Objective: Autosomal dominant hypercholesterolemias are due to defects in the LDL receptor (LDLR) gene, in the apolipoprotein B-100 gene or in the proprotein convertase subtilisin/kexin type 9 gene. The aim of this study was to identify and functionally characterize mutations in the LDLR gene that account for most cases of familial hypercholesterolemia (FH).
Methods: We enrolled 56 unrelated patients from Southern Italy with a clinical diagnosis of FH. The mutation screening was performed by direct sequencing of the promoter and the 18 exons of the LDLR gene and by multiplex ligation-dependent probe amplification (MLPA) analysis to search for large rearrangements.
Results and conclusion: We found 5 new mutations, the causative role of which was demonstrated by functional characterization performed by quantification of fluorescent LDL uptake in EBV-transformed B lymphocytes. These results enlarge the spectrum of FH-causative LDLR mutations. Lastly, screening for large rearrangements is highly recommended for the genetic diagnosis of FH.

1. Introduction

Autosomal dominant hypercholesterolemias are a heterogeneous group of genetic lipid disorders characterized by high levels of total and low-density lipoprotein (LDL) cholesterol. Causes of autosomal dominant hypercholesterolemias include defects in the LDL receptor (LDLR) gene, in its ligand apolipoprotein B-100 (ApoB), or in the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) gene [1]. The genetic defect remains unknown in a small percentage of FH patients [1].

More than 1200 mutations in the LDLR gene, at 19p13.1-p13.3 (MIM #606945), have been identified worldwide in FH patients (http://www.ucl.ac.uk/fh and http://www.umd.necker.fr); these include single-nucleotide mutations, duplications, deletions and splicing alterations [2].

The aim of this study was to identify and functionally characterize mutations in the LDLR gene in patients from Southern Italy with a clinical diagnosis of FH.

2. Materials and methods

2.1. Patients

Fifty-six unrelated patients from Southern Italy with clinically diagnosed FH were enrolled at the Dipartimento di Medicina Clinica e Sperimentale, Università degli Studi di Napoli Federico II. Relatives of 14/56 subjects were available. The FH diagnosis was based on the criteria established by the Società Italiana per lo Studio della Arteriosclerosi for the identification and treatment of dyslipidemias. The study was performed according to the current version of the Helsinki Declaration. Informed consent was obtained for each patient. The demographic and biochemical features of the patients are reported in Supplemental table.

* Corresponding author at: Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, via Pansini 5, 80131, Napoli, Italy. Tel.: +39 0817464200; fax: +39 0817462404.
E-mail address: fortunat@unina.it (G. Fortunato).

0021-9150/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.atherosclerosis.2009.11.051

2.2. Biochemical analysis

Biochemical markers were measured as described in Supplemental methods

2.3. Mutation screening

2.3.1. DNA sequencing of LDLR gene

Genomic DNA was isolated from whole blood EDTA samples using the Nucleon BACC2 kit (Amersham). The promoter and the 18 exons of the LDLR gene were amplified by PCR and directly sequenced. Details about PCR reactions and sequence analysis are reported in Supplemental methods. In the case of novel mutations, we verified: the absence in 150 chromosomes from normocholesterolemic individuals of the same ethnic group; segregation of the sequence variant in the relatives available; conservation of substituted aminoacid residues across homologous proteins by BLASTP.

2.3.2. Multiplex ligation-dependent probe analysis

We used multiplex ligation-dependent probe amplification (MLPA) to identify large rearrangements in the LDLR gene as described in Supplemental methods.

2.3.3. cDNA/RNA analysis

To identify splicing mutations, we performed RT-PCR analysis as described in Supplemental methods.

2.4. Mononuclear cell isolation and generation of Epstein-Barr Virus (EBV)-transformed cell lines

For each patient or control subject, peripheral blood mononuclear cells (PBMCs) were isolated from a blood sample as described in Supplemental methods. EBV-transformed cell lines were generated exposing PBMCs to free EBV particles, produced by an EBV-infected marmoset cell line (B95.8).

2.5. LDL uptake assay by fluorescent-activated cell sorter (FACS) flow cytometry

Residual LDLR activity was evaluated by measuring binding and uptake of a fluorescently labelled LDL as reported in Supplemental methods. Four samples from healthy subjects were used as reference.

Fluorescence intensities were measured with a FACSCanto (Becton Dickinson) flow cytometer according to the manufacturer's instructions. The results are expressed as the ratio between the median fluorescence intensity of cells from FH patients and the mean of the median fluorescence intensity of control cells. This ratio represents residual LDLR activity. All measurements have been performed in duplicate.

2.6. Statistical analysis

Statistical analyses were carried out with the Statistical Package for the Social Sciences version 16.0 (SPSS Inc., Chicago, IL, USA) as described in Supplemental methods.

3. Results

3.1. Mutation screening of the LDLR gene

We screened 56 unrelated FH patients from Southern Italy for mutations in the LDLR gene, and identified 16 known and 5 novel mutations in 43 subjects (Table 1). Direct sequencing of the promoter and encoding regions of the LDLR gene revealed mutations

Table 1
Mutations in the LDLR gene in patients with familial hypercholesterolemia from Southern Italy.

Location	Nucleotide change	Mutation type	Effect on protein[a]	Effect on protein (numeration without signal peptide)	Number of unrelated patients	Reference
Exon 2	c.116_117delGCinsAA	Nonsense	p.Cys39X	p.Cys18X	1	New
Exon 3	c.304C > T	Nonsense	p.Gln102X	p.Gln81X	2	[3]
Exon 4	c.352G > T	Missense	p.Asp118Tyr	p.Asp97Tyr	1	[4]
Exon 4	c.367T > C	Missense	p.Ser123Pro	p.Ser102Pro	1[b]	New
Exon 4	c.407A > T	Missense	p.Asp136Val	p.Asp115Val	1	New
Exon 4	c.542C > G	Missense	p.Pro181Arg	p.Pro160Arg	1	[5]
Exon 4	c.424_430delTCCTGCC	Deletion	p.Ser142ArgfsX62	p.Ser121ArgfsX62	1	New
Exon 8	c.1135T > C	Missense	p.Cys379Arg	p.Cys358Arg	2	[3]
Exon 9	c.1247G > C	Missense	p.Arg416Pro	p.Arg395Pro	1	[6]
Exon 9	c.1295T > C	Missense	p.Leu432Pro	p.Leu411Pro	1	[6]
Exon 10	c.1478_1479delCT	Deletion	p.Ser493CysfsX42	p.Ser472CysfsX42	1[b]	New
Exon 10	c.1567G > A	Missense	p.Val523Met	p.Val502Met	4	[7]
Exon 11	c.1646G > A	Missense	p.Gly549Asp	p.Gly528Asp	3	[8]
Exon 11	c.1698_1704delinsGCCCAAT	Deletion/insertion	p.Ile566_Leu568 delinsMetProAsn	p.Ile545_Leu547 delinsMetProAsn	1	[6]
Exon 12	c.1775G > A	Missense	p.Gly592Glu	p.Gly571Glu	5	[3]
Exon 14	c.2054C > T	Missense	p.Pro685Leu	p.Pro664Leu	1	[9]
Exon 15	c.2215C > T	Nonsense	p.Gln739X	p.Gln718X	1	[10]
Intron 3	c.313+1G > A	Splicing alteration	p.Leu64_Pro105delinsSer	p.Leu43_Pro84delinsSer	1	[11]
Intron 10	c.1586+1G > A	Splicing alteration	p.Thr454_Gly529del and p.Gly529_Phe530ins22	p.Thr433_Gly508del and p.Gly508_Phe509ins22	2	[4]
Intron 15	c.2312-3C > A	Splicing alteration	p.Ala771_796del	p.Ala750_775del	10	[6]
Exon 7	c.941-?_1060+?del	Gross deletion	p.Gly314_Glu353del	p.Gly293_Glu332del	1	[12]
Exons 11–12	c.1587-?_1845+?del	Gross deletion	p.Phe530ThrfsX49	p.Phe509ThrfsX49	1	[13]
Exons 13–15	c.1846-?_2311+?del	Gross deletion	p.Asp616LeufsX17	p.Asp595LeufsX17	1	[4]

[a] Mutations are cited according to Yamamoto's nomenclature derived from the UMD-LDLR database (www.umd.necker.fr).
[b] Mutations identified in a compound heterozygote.

Table 2
Measurement of residual LDLR activity by fluorescent-activated cell sorter (FACS) flow cytometry.

Sample	Median of fluorescence intensity	Residual activity of LDLR
Controls	1436[a]	100%
p.Ser123Pro and p.Ser493CysfsX42	459	32%
p.Asp136Val	1097	76%
p.Asp136Val and p.Gly592Glu	516	36%
p.Gly592Glu	1284	89%
p.Ser142ArgfsX62	683	49%
p.Ala771_796del	743	52%

[a] For controls the mean of median values of four samples were considered (standard deviation is 205).

in 40/56 patients (mutation rate 71.4%). In the 16 patients without detectable mutations at sequence analysis, we carried out an MLPA analysis to search for large rearrangements and found gross deletions in 3/16 subjects. Using these two methods, we identified mutations in 43/56 subjects (mutation rate, 76.8%).

The 5 new mutations were absent in 150 chromosomes from normal individuals. Three new mutations are a single nucleotide change (c.367T > C; c.407A > T; c.1295T > C) resulting in an aminoacid substitution (p.Ser123Pro; p.Asp136Val; p.Leu432Pro). The last two variations affect aminoacid residues conserved across the homologous LDL receptor proteins from 12 different examined species. Differently, p.Ser123Pro is conserved only in 7/12 species (chimpanzee, gorilla, orangutan, cat, rat, mouse and guinea pig) and it is not conserved in macaque, dog, pig, cow and zebrafish. Variation c.424_430delTCCTGCC causes a frameshift that resulted in the formation of a premature stop codon (p.Ser142ArgfsX62). Lastly, we identified a deletion/insertion mutation of two base pairs (c.116_117delGCinsAA) that gives rise to a premature stop codon producing a truncated protein of 38 aminoacids (p.Cys39X) that lacks most of the LDLR domains. Variants c.367T > C, c.407A > T, c.424_430delTCCTGCC and c.116_117delGCinsAA segregated with the FH phenotype, whereas c.1295T > C occurred in a single FH subject.

3.2. Functional characterization of LDLR

To determine the functional activity of the LDL receptor, we evaluated binding and uptake of fluorescent LDL (Dil-LDL) followed by FACS analysis.

Mutation c.367T > C (p.Ser123Pro) was identified in a compound heterozygote that also bears mutation c.1478_1479delCT (p.Ser493CysfsX42), which has already been identified [7] and characterized [4]. In this patient, the residual LDLR activity was about 32% compared to the normal controls (Table 2).

The residual LDLR activity in a cell line carrying mutation c.407A > T (p.Asp136Val) in heterozygosis was 76% (Table 2). During our screening, we found a member of the same family that was a compound heterozygote for both mutation c.407A > T (p.Asp136Val) and mutation c.1775G > A (p.Gly592Glu). In this compound heterozygote, the residual LDLR activity was 36% (Table 2). The residual LDLR activity in a patient carrying mutation c.1775G > A (p.Gly592Glu) in heterozygosis was 89% (Table 2).

Mutation c.424_430delTCCTGCC (p.Ser142ArgfsX62) was found in heterozygosis in 3 members of the same family. The residual LDLR activity evaluated in a cell line bearing this mutation was 49% (Table 2).

Finally, we carried out a functional characterization of the known mutation c.2312-3C > A that disrupts the donor splice site of intron 15 and results in a transcript that has an in-frame skipping of exon 16, as revealed by RT-PCR. The resulting protein has a deletion of a portion of the transmembrane domain (p.Ala771_796del). We measured the residual LDLR activity in 3 FH patients carrying c.2312-3C > A and obtained a mean value of 52% (range 50–53%) compared to normal controls (Table 2).

4. Discussion

In the attempt to enlarge the genetic spectrum of FH-causative mutations in the LDLR gene, we screened 56 unrelated subjects that had a clinical diagnosis of FH, and we found 5 novel mutations.

Regarding mutation c.367T > C (p.Ser123Pro), a variant (p.Ser123Thr) has previously been described in the same position [14] although this aminoacid residue is conserved only in homologue protein from 7/12 species. We characterized this mutation in a compound heterozygote for c.367T > C and c.1478_1479delCT (p.Ser493CysfsX42), and obtained a residual LDLR activity of 32%. This result indicates that, even if the patient carries the c.1478_1479delCT variant, which is more likely to induce the complete allele loss, mutation c.367T > C (p.Ser123Pro) is responsible for an additional decrease in the LDLR activity.

In our study, mutation c.407A > T (p.Asp136Val) was found alone and in compound heterozygosis together with c.1775G > A (p.Gly592Glu). Mutations c.407A > T and c.1775G > A had, in heterozygosis, a residual LDLR activity of 76% and 89% respectively. The presence of both mutations in a compound heterozygote resulted in residual LDLR activity of 36%. This seems to suggest that if a single mutation causes an appropriate compensatory increase of LDLR transcription, the compensatory increase of LDLR transcription in the presence of two mutated alleles is ineffective. The mutation c.1295T > C (p.Leu432Pro) occurs at the same codon as that previously described for the c.1294C > G (p.Leu432Val) mutation and it affects a highly conserved aminoacid located in the EGF precursor homologous domain [15].

Regarding mutation c.424_430delTCCTGCC (p.Ser142ArgfsX62), the residual LDLR activity was 49% indicating that the mutated allele has no functional activity.

Although we did not functionally characterize mutation c.116_117delGCinsAA (p.Cys39X), it might cause FH because it encodes a protein of only 38 aminoacids that probably has no functional activity for the mutated allele.

Despite the high genetic heterogeneity of FH-causing LDLR mutations, generally associated with a low frequency rate, in our population the c.2312-3C > A (p.Ala771_796del) variant accounts for 20% of all mutation cases. These data agree with previous reports and suggest this mutation has a founder effect in our region [6]. The mean residual LDLR activity of the c.2312-3C > A mutation was 52%, which implies that this mutation causes the complete functional loss of the protein encoded by the mutated allele. MLPA screening in our subjects without mutations at sequence analysis increased the mutation rate from 71.4% to 76.8%. This result supports the concept that genetic screening in FH disease should include the detection of large rearrangements.

In conclusion, our identification of 5 novel mutations in the LDLR gene enlarges the genetic heterogeneity of FH disease in our region and demonstrates the causative role of these mutations.

Acknowledgements

We thank Jean Ann Gilder for text revision and editing. We gratefully acknowledge the contribution of the Biobank of CEINGE for generation of EBV-transformed cell lines. Work supported by grants from CEINGE - Regione Campania (Protocollo di intesa-Del.G.R.29/12/2007), MUR art. 5.2 and Regione Campania L.R. no. 5/2005.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.atherosclerosis.2009.11.051.

References

[1] Varret M, Abifadel M, Rabès JP, Boileau C. Genetic heterogeneity of autosomal dominant hypercholesterolemia. Clin Genet 2008;73:1–13.
[2] Leigh SE, Foster AH, Whittall RA, Hubbart CS, Humphries SE. Update and analysis of the University College London low density lipoprotein receptor familial hypercholesterolemia database. Ann Hum Genet 2008;72:485–98.
[3] Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. Hum Mutat 1992;1:445–66.
[4] Bertolini S, Cassanelli S, Garuti R, et al. Analysis of LDL receptor gene mutations in Italian patients with homozygous familial hypercholesterolemia. Arterioscler Thromb Vasc Biol 1999;19:408–18.
[5] Chaves FJ, Real JT, García-García AB, et al. Genetic diagnosis of familial hypercholesterolemia in a South European outbreed population: influence of low-density lipoprotein (LDL) receptor gene mutations on treatment response to simvastatin in total, LDL, and high-density lipoprotein cholesterol. J Clin Endocrinol Metab 2001;86:4926–32.
[6] Liguori R, Bianco AM, Argiriou A, et al. LDL receptor cDNA sequence analysis in familial hypercholesterolemia patients: 5 novel mutations with high prevalence in families originating from Southern Italy. Hum Mutat 2001;17:433.
[7] Cavanaugh JA, Easteal S, Simons LA, Thomas DW, Serjeantson SW. FH-Sydney 1 and 2: two novel frameshift mutations in exon 10 of the low-density lipoprotein receptor gene detected by heteroduplex formation. Hum Mutat 1994;4:276–80.
[8] Hobbs HH, Russell DW, Brown MS, Goldstein JL. The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. Annu Rev Genet 1990;24:133–70.
[9] Soutar AK, Knight BL, Patel DD. Identification of a point mutation in growth factor repeat C of the low density lipoprotein-receptor gene in a patient with homozygous familial hypercholesterolemia that affects ligand binding and intracellular movement of receptors. Proc Natl Acad Sci USA 1989;86:4166–70.
[10] Yu W, Nohara A, Higashikata T, et al. Molecular genetic analysis of familial hypercholesterolemia: spectrum and regional difference of LDL receptor gene mutations in Japanese population. Atherosclerosis 2002;165:335–42.
[11] Leren TP, Solberg K, Rødningen OK, Tonstad S, Ose L. Two founder mutations in the LDL receptor gene in Norwegian familial hypercholesterolemia subjects. Atherosclerosis 1994;111:175–82.
[12] Górski B, Kubalska J, Naruszewicz M, Lubiński J. LDL-R and Apo-B-100 gene mutations in Polish familial hypercholesterolemias. Hum Genet 1998;102:562–5.
[13] Bertolini S, Garuti R, Lelli W, et al. Four novel partial deletions of LDL-receptor gene in Italian patients with familial hypercholesterolemia. Arterioscler Thromb Vasc Biol 1995;15:81–8.
[14] Mihaylov VA, Horvath AD, Savov AS, et al. Screening for point mutations in the LDL receptor gene in Bulgarian patients with severe hypercholesterolemia. J Hum Genet 2004;49:173–6.
[15] Ebhardt M, Schmidt H, Doerk T, et al. Mutation analysis in 46 German families with familial hypercholesterolemia: identification of 8 new mutations. Hum Mutat 1999;13:257.

methods

An improved method on stimulated T-lymphocytes to functionally characterize novel and known LDLR mutations[S]

Maria Romano,*,† Maria Donata Di Taranto,§ Peppino Mirabelli,*,† Maria Nicoletta D'Agostino,*,† Arcangelo Iannuzzi, Gennaro Marotta,†† Marco Gentile,†† Maddalena Raia,* Rosa Di Noto,*,† Luigi Del Vecchio,*,† Paolo Rubba,†† and Giuliana Fortunato[1],*,†**

CEINGE - Biotecnologie Avanzate,* Napoli, Italy; Dipartimento di Biochimica e Biotecnologie Mediche,† and Dipartimento di Medicina Clinica e Sperimentale,†† Università degli Studi di Napoli Federico II, Napoli, Italy; Istituto di Ricovero e Cura a Carattere Scientifico Fondazione SDN,§ Napoli, Italy; and Centro per la diagnosi e la terapia delle dislipidemie familiari, Azienda Ospedaliera Antonio Cardarelli,** Napoli, Italy

Abstract The main causes of familial hypercholesterolemia (FH) are mutations in LDL receptor (LDLR) gene. Functional studies are necessary to demonstrate the LDLR function impairment caused by mutations and would be useful as a diagnostic tool if they allow discrimination between FH patients and controls. In order to identify the best method to detect LDLR activity, we compared continuous Epstein-Barr virus (EBV)-transformed B-lymphocytes and mitogen stimulated T-lymphocytes. In addition, we characterized both novel and known mutations in the LDLR gene. T-lymphocytes and EBV-transformed B-lymphocytes were obtained from peripheral blood of 24 FH patients and 24 control subjects. Functional assays were performed by incubation with fluorescent LDL followed by flow cytometry analysis. Residual LDLR activity was calculated normalizing fluorescence for the mean fluorescence of controls. With stimulated T-lymphocytes we obtained a better discrimination capacity between controls and FH patients compared with EBV-transformed B-lymphocytes as demonstrated by receiver operating characteristic (ROC) curve analysis (the areas under the curve are 1.000 and 0.984 respectively; $P < 0.0001$ both). The characterization of LDLR activity through T-lymphocytes is more simple and faster than the use of EBV-transformed B-lymphocytes and allows a complete discrimination between controls and FH patients. Therefore the evaluation of residual LDLR activity could be helpful not only for mutation characterization but also for diagnostic purposes.—Romano,M., M. Donata Di Taranto, P. Mirabelli, M. N. D'Agostino, A. Iannuzzi, G. Marotta, M. Gentile, M. Raia, R. Di Noto L. Del Vecchio, P. Rubba, and G. Fortunato. **An improved method on stimulated T-lymphocytes to functionally characterize novel and known LDLR mutations. *J. Lipid Res.* 2011.** 52: **2095–2100.**

Supplementary key words LDL receptor • familial hypercholesterolemia • EBV-transformed B-lymphocytes • mitogen stimulated T-lymphocytes • functional activity

Familial hypercholesterolemia (FH) is an autosomal dominant disorder characterized by elevated plasma LDL cholesterol, tendon xanthomas, and premature coronary heart disease. FH is mostly caused by mutations within the low density lipoprotein receptor (LDLR; MIM# 143890) gene, or in its ligand apoB-100 (MIM# 107730), or in the proprotein convertase subtilisin/kexin type 9 (PCSK9; MIM# 607786) gene (1).

At present, more than 1,100 variants of LDLR gene have been listed in the LDLR databases underlying a high genetic heterogeneity of LDLR mutations. These are distributed across the 18 exons, introns, and the promoter region of the LDLR gene and include point mutations, insertions, deletions, and major rearrangements (2).

Although LDLR defects are primarily identified by genetic methods, it is extremely important to demonstrate a deficiency in the LDLR function of FH suspect patients through functional studies. Receptor assays reported so far include measurement of radiolabeled-LDL or fluorescently-labeled LDL binding and/or uptake in skin fibroblasts

Abbreviations: AUC, area under the receiver operating characteristic curves; CI, confidence interval; CV%, coefficient of variation; DiI-LDL, 1,19-dioctadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate-conjugated LDL; EBV, Epstein-Barr virus; FH, familial hypercholesterolemia; IMDM, Iscove's Modified Dulbecco's Medium; LDLR, LDL receptor; LPDS, lipoprotein deprived serum; PBMC, peripheral blood mononuclear cell; PCSK9, proprotein convertase subtilisin/kexin type 9; PHA, phytohemagglutinin; PMA, phorbol 12-myristate 13-acetate; ROC, receiver operating characteristic.

[1]To whom correspondence should be addressed.
e-mail: fortunat@unina.it

[S] The online version of this article (available at http://www.jlr.org) contains supplementary data in the form of one table and three figures.

This work was supported by grants from CEINGE Convenzione Regione Campania, DGRC 1901/2009 and from IRCCS Fondazione SDN.

Manuscript received 8 June 2011 and in revised form 8 August 2011.

Published, JLR Papers in Press, August 22, 2011
DOI 10.1194/jlr.D017772

(3, 4) or leukocytes (5, 6). The Epstein-Barr virus (EBV)-transformed B-lymphocytes show high LDLR levels and allow an unlimited supply of cells without being influenced by the patient's diet and drug treatment (7). The main problem of these functional assays was the bad separation between residual LDLR activity from FH patients and controls. The separation of different leukocyte populations obtains more accurate results (8); the selection of viable lymphocytes by stimulation with a mitogen or by flow cytometry gating improves the discrimination between patients and controls although it does not allow complete discrimination of FH heterozygote patients from healthy controls (9, 10).

The aim of this study was to improve the method of T-lymphocyte stimulation for LDLR activity evaluation through the comparison of different mitogen combinations. We also aimed to compare continuous EBV-transformed-lymphocytes and mitogen stimulated T-lymphocytes obtained from peripheral blood samples in order to identify the best method to detect LDLR activity by fluorescently-labeled LDL. In addition, we characterized both novel and known mutations in the LDLR gene.

MATERIALS AND METHODS

FH patients and control subjects

Patients with clinically diagnosed FH were enrolled at the Dipartimento di Medicina Clinica e Sperimentale, Università degli Studi di Napoli Federico II and at the AORN Cardarelli, Napoli. The FH diagnosis was based on the criteria established by the Società Italiana per lo Studio della Arteriosclerosi for the identification and treatment of dyslipidemias. Twenty-four patients bearing new and known mutations were selected for functional characterization. As healthy controls, 24 unrelated subjects were selected from voluntary donors of the Centro Trasfusionale of the AORN Cardarelli, Napoli, on the basis of normal lipid levels and absence of LDLR mutations. The supplementary table reports characteristics of FH patients and controls. The study was performed according to the current version of the Helsinki Declaration. Informed consent was obtained for each patient.

Mutation screening

Molecular analysis for the identification of LDLR mutations was performed as previously described (11). Briefly, the promoter and the 18 exons of the LDLR gene were amplified by PCR and directly sequenced. In the case of novel mutations, we verified the absence in 150 chromosomes from normocholesterolemic individuals of the same ethnic group, the segregation of the sequence variant in the relatives available, and the conservation of substituted amino acid residues across homologous proteins by BLASTP.

Mononuclear cell isolation and generation of EBV-transformed B-lymphocytes

Peripheral blood mononuclear cells (PBMCs) were isolated from 9 ml of EDTA blood collected from each subject enrolled in this study. The sample was diluted 1 in 2 with PBS without Ca^{2+} and Mg^{2+} pH 7.4, and then layered on Ficoll-Paque (Lymphocyte Separation Medium, MP Biomedicals). After centrifugation at 1,500 rpm for 30 min at 4°C, mononuclear cells were recovered at interface, washed three times in PBS and stored in 90% FBS (Lonza) and 10% DMSO in liquid nitrogen until required for use.

EBV-transformed cell lines were generated by exposing PBMCs to free EBV particles, produced by an EBV-infected marmoset cell line (B95.8). After EBV infection, transformed PBMCs were cultured in an incubator at 37°C in a 5% CO_2 atmosphere in Iscove's Modified Dulbecco's Medium (IMDM, Sigma-Aldrich) supplemented with 20% FBS and 2% Ultraglutamine (Lonza). The continuous cell lines bearing the LDLR mutations were also analyzed to verify the presence of DNA variations. In order to upregulate LDLR, the cells were incubated in a medium containing a lipoprotein deprived serum (LPDS, Sigma-Aldrich) for 48 h before being analyzed.

Peripheral blood T-lymphocytes stimulation

After two wash steps in FBS, thawed PBMC were counted and seeded in a 24-well plate (BD-Falcon) at 1×10^6 cells/ml in IMDM supplemented with 10% human LPDS to induce LDLR upregulation and 2% Ultraglutamine. Stimulation of T-lymphocytes was carried out for 48 h at 37°C and 5% CO_2 with three different combinations of mitogens: *1*) 5μg/ml phytohemagglutinin (PHA) (Sigma-Aldrich), *2*) PHA and 10 ng/ml phorbol 12-myristate 13-acetate (PMA, Sigma-Aldrich), and *3*) 1 μM ionomycin (Sigma-Aldrich) and 10 ng/ml PMA according to Makar et al. (12).

LDL uptake assay by fluorescent-activated cell sorter flow cytometry

Residual LDLR activity was evaluated by measuring the binding and uptake of a fluorescently labeled LDL, namely 1,19-dioctadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate-conjugated LDL (DiI-LDL, Invitrogen), in EBV-transformed B lymphocytes and mitogen-stimulated T-lymphocytes. Briefly, cells were collected and diluted at 1×10^6 cells/ml in DMEM:F12 (Sigma-Aldrich) without phenol red and LPDS. Two aliquots of cells for each sample analyzed were then incubated for 3 h at 37°C and 5% CO_2 as follows: *a*) with 30 mg/L unlabeled LDL (Sigma-Aldrich), to measure background fluorescence; and *b*) with 10 mg/L DiI-LDL to estimate the maximum LDL binding and uptake. In preliminary experiments we performed an additional incubation with 30 mg/L unlabeled LDL plus 10 mg/L DiI-LDL (ratio 3:1) to evaluate the displacement of DiI-LDL and in this condition, we observed a reduction of cell fluorescence indicating the specific binding of DiI-LDL (supplementary Fig. I). After three washes in with PBS without Ca^{2+} and Mg^{2+}, fluorescence intensities were measured with a FACSCanto (Becton-Dickinson) flow cytometer according to the manufacturer's instructions. Forward scatter and side scatter gates were established to exclude dead cells and cell debris. An example of flow cytometry gating is shown in supplementary Fig. II that also shows the percentage of $CD3^+$ cells (T-lymphocytes) obtained with the ionomycin plus PMA stimulation. For each experiment, 10,000 events were counted.

The results are expressed as the median of fluorescence and as the ratio between the median fluorescence intensity of cells from FH patients and the mean of the median fluorescence intensity of control cells. This ratio represents residual LDLR activity. All measurements were performed in duplicate.

Statistical analysis

The normality of variable distributions was evaluated with Kolmogorov-Smirnov test. Continuous variables were indicated as mean ± SD. Differences between the continuous variables were evaluated with *t*-test. A value of $P < 0.05$ was considered statistically significant. Statistical analyses were carried out using the

statistical Predictive Analytics Software version 18.0 (SPSS Inc.). Analysis of receiver operating characteristic (ROC) curves and dot diagrams was performed with MedCalc Version 11.5.1. The statistical significance of the area under the ROC curves (AUC) was calculated against the null hypothesis AUC = 0.5 as recommended by DeLong et al. (13). The threshold values were determined by the farthest point from the bisector of the ROC curve.

RESULTS

Mutation screening

We identified four new mutations absent in 150 chromosomes from normal individuals (**Table 1**). These alterations are novel missense mutations (c.440C>T; c.974G>A; c.1130G>T; c.1739C>T) that result in a single amino acid substitution in LDL receptor protein (p.Thr147Ile; p.Cys325Tyr; p.Cys377Phe; p.Ser580Phe). The novel mutations cause changes of amino acid residues conserved across fourteen species examined during our study.

T-lymphocytes mitogen selection

In order to select the best way to stimulate T-lymphocytes able to express LDLR, we compared three known procedures for in vitro growth stimulation of T-cells. The flow cytometry analysis showed that the amount of viable T-lymphocytes in cells treated with *1*) PHA, *2*) PHA plus PMA, and *3*) ionomycin plus PMA was 76.4%, 83.8%, and 87.3%, respectively. As regards LDL-R expression evaluated by the uptake of DiI-LDL, in case of PHA stimulation, the uptake of DiI-LDL was lower than that observed with PHA plus PMA and ionomycin plus PMA (supplementary Fig. I). Due to the higher yield of live T-lymphocytes obtained with ionomycin plus PMA with respect to PHA plus PMA and considering the LDLR upregulation in T-lymphocytes described by Makar et al. (12) using ionomycin plus PMA, we selected this method for further experiments.

TABLE 1. Medians of fluorescent intensity and LDLR residual activities evaluated on EBV-transformed B-lymphocytes and stimulated T-lymphocytes from control subjects.

Control ID	Median of fluorescent intensity on EBV-transformed B-lymphocytes	Median of fluorescent intensity on stimulated T-lymphocytes	LDLR residual activity on EBV-transformed B-lymphocytes	LDLR residual activity on stimulated T-lymphocytes
Control 1	2,173	1,132	124.0	111.9
Control 2	1,654	1,161	94.4	114.8
Control 3	1,754	1,349	100.1	133.4
Control 4	2,304	931	131.5	92.0
Control 5	1,981	1,106	113.0	109.3
Control 6	2,150	928	122.7	91.7
Control 7	2,103	1,103	120.0	109.0
Control 8	1,524	948	86.9	93.7
Control 9	1,289	869	73.5	85.9
Control 10	2,056	952	117.3	94.1
Control 11	1,924	904	109.8	89.4
Control 12	1,257	799	71.7	79.0
Control 13	1,582	–	90.2	–
Control 14	1,387	–	79.1	–
Control 15	1,238	–	70.6	–
Control 16	1,658	–	94.6	–
Control 17	–	1,048	–	103.6
Control 18	–	963	–	95.2
Control 19	–	905	–	89.5
Control 20	–	1,231	–	121.7
Control 21	–	1,101	–	108.9
Control 22	–	940	–	92.9
Control 23	–	882	–	87.2
Control 24	–	960	–	94.9
Mean ± SD	1,752 ± 355	1,011 ± 138	100.0 ± 20.3	100.0 ± 13.6
CV%	20.3	13.6	20.3	13.6

Missing data are relative to unavailable samples. CV%, coefficient of variation.

Functional characterization of LDLR activity

The median of fluorescence intensity and the residual LDLR activity evaluated in EBV-transformed B-lymphocytes and in ionomycin plus PMA stimulated T-lymphocytes from healthy subjects with normal lipid profile and no mutations in LDLR are reported in Table 1. A representative set of data from flow cytometry analysis is shown in supplementary Fig. III. The residual LDLR activity was calculated using the mean of median values of all analyzed controls as areference (1,752 for EBV-transformed B-lymphocytes and 1,011 for stimulated T-lymphocytes). As for controls, this calculation represents a data normalization allowing the comparison between the two methods used. The minimum and maximum values of residual LDLR activity evaluated on EBV-transformed B-lymphocytes from controls were 71% and 132%, respectively, whereas the minimum and maximum values of residual LDLR activity evaluated on stimulated T-lymphocytes were 79% and 133%, respectively. The median of fluorescence intensity and the residual LDLR activity in EBV-transformed B-lymphocytes and in stimulated T-lymphocytes from FH patients with mutations in LDLR gene are reported in **Table 2**. The residual LDLR activities were statistically different between control subjects and FH patients analyzed by both methodologies ($P < 0.0001$ both). Regarding the functional characterization of new mutations, we observed low residual LDLR activities consistent with a heterozygous status of FH (Table 2).

Figure 1 shows a graphical representation of residual LDLR activity of each control and FH patient evaluated by the two methods. We observed that, using EBV-transformed B-lymphocytes, there is an overlap of residual LDLR activities between controls and patients in the range 71–75%, whereas, using stimulated T-lymphocytes, no value in the range 65–79% was observed for controls nor for FH patients. The use of stimulated T-lymphocytes allows a complete discrimination between controls and FH patients.

Analysis of ROC curves

The ROC curves are constructed using the residual LDLR activity evaluated in EBV-transformed B-lymphocytes and in stimulated T-lymphocytes as a discriminator parameter between FH patients and healthy controls (**Fig. 2**). The AUC of the methods based on EBV-transformed B-lymphocytes and on stimulated T-lymphocytes are 0.984 [95% confidence interval (CI) 0.880–1.000] and 1.000 (95% CI 0.907–1.000), respectively ($P < 0.0001$ both). The comparison of the AUC of the two methods does not reveal any statistically significant difference ($P = 0.41$).

TABLE 2. Medians of fluorescent intensity and residual LDLR activities evaluated on EBV-transformed B-lymphocytes and stimulated T-lymphocytes from FH patients

Nucleotide change	Effect on protein	Reference	Status	Median of fluorescent intensity on EBV-transformed B-lymphocytes	Median of fluorescent intensity on stimulated T-lymphocytes	LDLR residual activity on EBV-transformed B-lymphocytes	LDLR residual activity on stimulated T-lymphocytes
c.116_117 delGCinsAA	p.Cys39X	11	heterozygote	1,113	444	63.5	43.9
c.367T>C and c.1478_1479delCT	p.Ser123Pro and p.Ser493CysfsX42	11 and 14	compound heterozygote	459	–	26.1	–
c.407A>T	p.Asp136Val	11	heterozygote	1,097	–	62.6	–
c.407A>T and c.1775G>A	p.Asp136Val and p.Gly592Glu	11 and 15	compound heterozygote	516	220	29.4	21.7
c.424_430 delTCCTGCC	p.Ser142ArgfsX62	11	heterozygote	1,186	647	67.6	64.0
c.440C>T	p.Thr147Ile	new	heterozygote	982	–	56.0	–
c.440C>T	p.Thr147Ile	new	heterozygote	1,241	562	70.8	55.6
c.974G>A	p.Cys325Tyr	new	heterozygote	949	572	54.1	56.6
c.1130G>T	p.Cys377Phe	new	heterozygote	750	–	42.8	–
c.1135T>C	p.Cys379Arg	15	heterozygote	1,157	566	66.0	56.0
c.1187-10G>A	p.Gly396AspfsX19	16	heterozygote	944	–	53.8	–
c.1187-10G>A	p.Gly396AspfsX19	16	heterozygote	1,317	615	75.1	60.8
c.1567G>A	p.Val523Met	17	heterozygote	635	607	36.2	60.0
c.1646G>A	p.Gly549Asp	18	heterozygote	979	354	55.8	35.0
c.1646G>A	p.Gly549Asp	18	heterozygote	831	653	47.4	64.6
c.1646G>A	p.Gly549Asp	18	heterozygote	1,149	503	65.5	49.7
c.1739C>T	p.Ser580Phe	new	heterozygote	1,152	646	65.7	63.9
c.1739C>T and c.1646G>A	p.Ser580Phe and p.Gly549Asp	new and 18	compound heterozygote	741	306	42.2	30.2
c.1775G>A	p.Gly592Glu	15	heterozygote	1,284	–	73.2	–
c.1775G>A	p.Gly592Glu	15	homozygote	682	537	38.9	53.1
c.1846-?_2311+?del	p.Asp616LeufsX17	17	heterozygote	338	367	19.2	36.3
c.2312-3C>A	p.Ala771_796del	19	heterozygote	724	567	41.3	56.1
c.2312-3C>A	p.Ala771_796del	19	heterozygote	763	478	43.5	47.2
c.2312-3C>A	p.Ala771_796del	19	heterozygote	–	653	–	64.6

Missing data are relative to unavailable samples.

By the analysis of ROC curves, the suggested threshold of residual LDLR activity for discrimination between controls and patients is 67.7% for EBV-transformed B-lymphocytes (86.96% sensitivity and 100% specificity) and 64.6% for stimulated T-lymphocytes (100% sensitivity and 100% specificity).

DISCUSSION

The functional characterization of LDLR mutations by use of EBV-transformed B-lymphocytes provides the advantages of continuous cell lines that undergo rapid proliferation combined with a long lifespan. These are easily reestablished in culture after being frozen and are useful in the creation of cell banks for future molecular studies. These advantages enable a reliable characterization of novel and known mutations in LDLR gene and other genetic disorders (20). However, EBV-transformed B-lymphocytes must be used carefully because of decreased genetic stability and possible changes in cellular processes induced by viral transformation (20). The generation of EBV-transformed B-lymphocytes is wasteful, time-consuming, and often results tare unproductive depending on the quantity of leukocytes isolated from each subject.

In order to overcome these problems, we have improved an assay to evaluate LDLR activity in peripheral-blood lymphocytes using ionomycin plus PMA-induced proliferating T-lymphocytes. We selected this mitogen combination because in our experimental conditions, it gives the best results. In other studies, functional assays were conducted on T-lymphocytes using only PHA (21, 22) although the comparison between different mitogens was not performed. For the first time, we compared two cell culture techniques useful in detecting residual LDLR activity in FH patients: EBV-transformed B-lymphocytes and mitogen stimulated T-lymphocytes both treated with fluorescently-labeled LDL. In controls, we observed a higher median of fluorescent intensity in EBV-transformed B-lymphocytes than in stimulated T-lymphocytes, suggesting the highest LDLR expression in the first cell type. The residual LDLR activity evaluated on stimulated T-lymphocytes of controls shows a more limited range of values, suggesting a minor LDLR expression variability.

The minimum value of residual LDLR activity in controls is higher for stimulated T-lymphocytes than for EBV-transformed B-lymphocytes (79% and 71%, respectively) suggesting that the T-lymphocytes based method has a better capacity to detect functional activity than the other one. With ionomycin plus PMA stimulated T-lymphocytes, no value was observed in the range 65–79% (discriminating range) for controls nor for FH patients, whereas using EBV-transformed B-lymphocytes, we observed an overlap in residual LDLR activities between controls and patients in the range 71–75% (absence of a discriminating range). Then, using stimulated T-lymphocytes, we obtained a complete discrimination between FH patients and normocholesterolemic

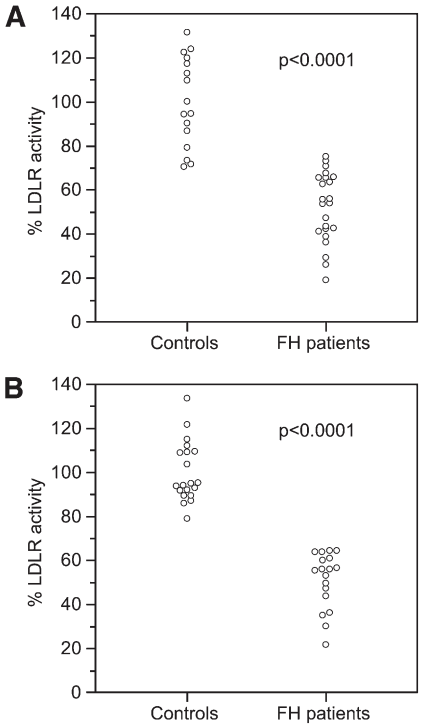

Fig. 1. LDLR activity in controls and FH patients evaluated in EBV-transformed B-lymphocytes and stimulated T-lymphocytes. A: EBV-transformed B-lymphocytes; B: stimulated T-lymphocytes. Data are presented as percentage of LDLR activity. The p-values are related to the comparison between controls and FH patients.

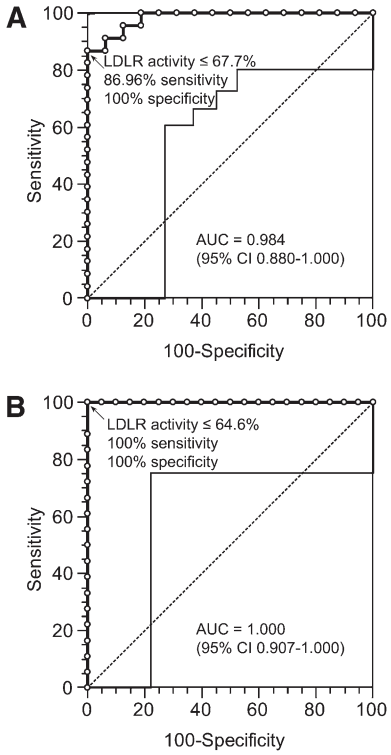

Fig. 2. Receiver operating characteristic (ROC) curves of LDLR activity evaluated in EBV-transformed B-lymphocytes and stimulated T-lymphocytes. A: EBV-transformed B-lymphocytes; B: stimulated T-lymphocytes. The ROC curve is indicated with bold line and open circles represent the criterion points. Light line indicate the 95% confidence interval (CI). Dotted line indicate the bisector. AUC, area under the curve.

subjects, although we observed reduced LDLR activities in FH patients with both methods, which is in agreement with the presence of LDLR mutations. The technical advancement reached with our method consists in the complete discrimination between FH patients and controls, a feature that could allow use of this functional assay for screening purposes.

The analysis of ROC curves in stimulated T-lymphocytes suggests a threshold value of residual LDLR activity equal to 64.6% with a 100% sensitivity and a 100% specificity, although the use of a higher threshold in the discriminating range (65–79%) could improve the diagnostic sensitivity of the method.

To date, many techniques have been used to diagnose FH disease, including serum lipid profile and molecular methods such as DNA sequencing for the screening of point mutations, multiplex ligation-dependent probe amplification to identify large rearrangements, and RNA analysis for splicing variants (2, 23–25). The residual LDLR activity evaluated on stimulated T-lymphocytes (ionomycin plus PMA used as mitogen stimulators) may be used as a screening method for identification of LDLR defects in patients with a clinical suspect of FH. In addition, our method is relatively easy to handle and appears less laborious and time-consuming than other described techniques such as the use of EBV-transformed B-lymphocytes whose generation requires to work in class 1 biosafety level laboratories.

In conclusion, our method based on the DiI-LDL uptake in ionomycin plus PMA-stimulated T-lymphocytes is simple and fast and improves the detection of LDL receptor activity defects, allowing complete discrimination between controls and FH patients. Furthermore, the evaluation of residual LDLR activity could be helpful not only for mutation characterization but also for diagnostic purposes.

The authors gratefully acknowledge the contribution of the Biobank of CEINGE for isolation of peripheral blood mononuclear cells and generation of EBV-transformed cell lines.

REFERENCES

1. Soutar, A. K. 2010. Rare genetic causes of autosomal dominant or recessive hypercholesterolaemia. *IUBMB Life.* **62:** 125–131.
2. Varret, M., M. Abifadel, J. P. Rabès, and C. Boileau. 2008. Genetic heterogeneity of autosomal dominant hypercholesterolemia. *Clin. Genet.* **73:** 1–13.
3. Goldstein, J. L., and M. S. Brown. 1974. Binding and degradation of low density lipoproteins by cultured human fibroblasts. Comparison of cells from a normal subject and from a patient with homozygous familial hypercholesterolemia. *J. Biol. Chem.* **249:** 5153–5162.
4. Leitersdorf, E., H. H. Hobbs, A. M. Fourie, M. Jacobs, D. R. van der Westhuyzen, and G. A. Coetzee. 1988. Deletion in the first cysteine-rich repeat of low density lipoprotein receptor impairs its transport but not lipoprotein binding in fibroblasts from a subject with familial hypercholesterolemia. *Proc. Natl. Acad. Sci. USA.* **85:** 7912–7916.

5. Schmitz, G., T. Brüning, E. Kovacs, and S. Barlage. 1993. Fluorescence flow cytometry of human leukocytes in the detection of LDL receptor defects in the differential diagnosis of hypercholesterolemia. *Arterioscler. Thromb.* **13:** 1053–1065.
6. Løhne, K., P. Urdal, T. P. Leren, S. Tonstad, and L. Ose. 1995. Standardization of a flow cytometric method for measurement of low-density lipoprotein receptor activity on blood mononuclear cells. *Cytometry.* **20:** 290–295.
7. Chan, P., C. Jones, R. Lafrenière, and H. G. Parsons. 1997. Surface expression of low density lipoprotein receptor in EBV-transformed lymphocytes: characterization and use for studying familial hypercholesterolemia. *Atherosclerosis.* **131:** 149–160.
8. Raungaard, B., F. Heath, J. U. Brorholt-Petersen, H. K. Jensen, and O. Faergeman. 1999. Flow cytometric assessment of LDL receptor activity in peripheral blood mononuclear cells compared to gene mutation detection in diagnosis of heterozygous familial hypercholesterolemia. *Cytometry.* **36:** 52–59.
9. Tada, H., M. Kawashiri, T. Noguchi, M. Mori, M. Tsuchida, M. Takata, A. Nohara, A. Inazu, J. Kobayashi, A. Yachie, et al. 2009. A novel method for determining functional LDL receptor activity in familial hypercholesterolemia: application of the CD3/CD28 assay in lymphocytes. *Clin. Chim. Acta.* **400:** 42–47.
10. Urdal, P., T. P. Leren, S. Tonstad, P. K. Lund, and L. Ose. 1997. Flow cytometric measurement of low density lipoprotein receptor activity validated by DNA analysis in diagnosing heterozygous familial hypercholesterolemia. *Cytometry.* **30:** 264–268.
11. Romano, M., M. D. Di Taranto, M. N. D'Agostino, G. Marotta, M. Gentile, G. Abate, P. Mirabelli, R. Di Noto, L. Del Vecchio, P. Rubba, et al. 2010. Identification and functional characterization of LDLR mutations in familial hypercholesterolemia patients from Southern Italy. *Atherosclerosis.* **210:** 493–496.
12. Makar, R. S., P. E. Lipsky, and J. A. Cuthbert. 1994. Non-sterol regulation of low density lipoprotein receptor gene expression in T cells. *J. Lipid Res.* **35:** 1888–1895.
13. DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* **44:** 837–845.
14. Cavanaugh, J. A., S. Easteal, L. A. Simons, D. W. Thomas, and S. W. Serjeantson. 1994. FH-Sydney 1 and 2: two novel frameshift mutations in exon 10 of the low-density lipoprotein receptor gene detected by heteroduplex formation. *Hum. Mutat.* **4:** 276–280.
15. Hobbs, H. H., M. S. Brown, and J. L. Goldstein. 1992. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum. Mutat.* **1:** 445–466.
16. Wang, J., E. Huff, L. Janecka, and R. A. Hegele. 2001. Low density lipoprotein receptor (LDLR) gene mutations in Canadian subjects with familial hypercholesterolemia, but not of French descent. *Hum. Mutat.* **18:** 359.
17. Hobbs, H. H., D. W. Russell, M. S. Brown, and J. L. Goldstein. 1990. The LDL Receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. *Annu. Rev. Genet.* **24:** 133–170.
18. Bertolini, S., S. Cassanelli, R. Garuti, M. Ghisellini, M. L. Simone, M. Rolleri, P. Masturzo, and S. Calandra. 1999. Analysis of LDL receptor gene mutations in Italian patients with homozygous familial hypercholesterolemia. *Arterioscler. Thromb. Vasc. Biol.* **19:** 408–418.
19. Liguori, R., A. M. Bianco, A. Argiriou, P. Pauciullo, A. Giannino, P. Rubba, and V. De Simone. 2001. LDL receptor cDNA sequence analysis in familial hypercholesterolemia patients: 5 novel mutations with high prevalence in families originating from southern Italy. *Hum. Mutat.* **17:** 433.
20. Sie, L., S. Loong, and E. K. Tan. 2009. Utility of lymphoblastoid cell lines. *J. Neurosci. Res.* **87:** 1953–1959.
21. Chan, P. C., A. Edwards, R. Lafrenière, and H. G. Parsons. 1998. Improved detection of familial hypercholesterolemia by determining low density lipoprotein receptor expression in mitogen-induced proliferating lymphocytes. *J. Lipid Res.* **39:** 2261–2270.
22. Cuthbert, J. A., C. A. East, D. W. Bilheimer, and P. E. Lipsky. 1986. Detection of familial hypercholesterolemia by assaying functional low-density-lipoprotein receptors on lymphocytes. *N. Engl. J. Med.* **314:** 879–883.
23. Minhas, R., S. E. Humphries, N. Qureshi, and H. A. Neil; NICE Guideline Development Group. 2009. Controversies in familial hypercholesterolaemia: recommendations of the NICE Guideline Development Group for the identification and management of familial hypercholesterolaemia. *Heart.* **95:** 584–587.
24. Taylor, A., D. Wang, K. Patel, R. Whittall, G. Wood, M. Farrer, R. D. Neely, S. Fairgrieve, D. Nair, M. Barbir, et al. 2010. Mutation detection rate and spectrum in familial hypercholesterolaemia patients in the UK pilot cascade project. *Clin. Genet.* **77:** 572–580.
25. Holla, Ø. L., S. Nakken, M. Mattingsdal, T. Ranheim, K. E. Berge, J. C. Defesche, and T. P. Leren. 2009. Effects of intronic mutations in the LDLR gene on pre-mRNA splicing: comparison of wet-lab and bioinformatics analyses. *Mol. Genet. Metab.* **96:** 245–252.

Investigation of Single Nucleotide Polymorphisms associated to Familial Combined Hyperlipidemia with Random Forests

Antonino Staiano[1], Maria Donata Di Taranto[2], Elena Bloise[3], Maria Nicoletta D'Agostino[3,4], Antonietta D'Angelo[3,4], Gennaro Marotta[5], Marco Gentile[5], Fabrizio Jossa[5], Arcangelo Iannuzzi[6], Paolo Rubba[5], and Giuliana Fortunato[3,4]

[1] Dipartimento di Scienze Applicate, Università di Napoli "Parthenope", Isola C4, Centro Direzionale, 80143, Napoli, Italy
antonino.staiano@uniparthenope.it
[2] IRCCS Fondazione SDN, Napoli, Italy
[3] CEINGE S.C.a r.l. Biotecnologie Avanzate, Napoli, Italy
[4] Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, Napoli, Italy
[5] Dipartimento di Medicina Clinica e Sperimentale, Università degli Studi di Napoli Federico II, Napoli, Italy
[6] Centro per la diagnosi e la terapia delle dislipidemie familiari, Azienda Ospedaliera A. Cardarelli, Napoli, Italy

Abstract. Single nucleotide polymorphisms (SNPs) are the foremost part of many genome association studies. Selecting a subset of SNPs that is sufficiently informative but still small enough to reduce the genotyping overhead is an important step towards disease-gene association. In this work, a Random Forest (RF) approach to informative SNPs selection in Familial Combined Hyperlipidemia (FCH) is proposed. FCH is the most common form of familial hyperlipidemia. Affected patients have elevated levels of plasma triglycerides and/or total cholesterol and show increased risk of premature coronary heart disease. In order to identify susceptibility markers for FCH we perform the analysis of 21 SNPs in ten genes associated with high cardiovascular risk. RF appears to be a useful technique in identifying gene polymorphisms involved in FCH: the identified SNPs confirmed some variants in a couple of genes as genetic markers of FCH as proved in various studies in scientific literature and lead us to report for the first time a further gene association with FCH. This result could be promising and encourages to further investigate on the role of the identified gene in the development of FCH phenotype.

Keywords: Machine learning, feature selection, random forests, genome association, single nucleotide polymorphisms, familial combined hyperlipidemia

1 Introduction

Genome and genome-wide association studies (GA) have revolutionized genetics making it possible the search for genetic factors that influence common complex

traits and the characterization of the effects of those factors [5]. Single nucleotide polymorphisms (SNPs) are at the forefront of many disease-gene association studies. SNPs are mutations at a single nucleotide position that occurred during evolution and were passed on through heredity, accounting for most of the genetic variation among different individuals. The number of SNPs has been estimated to be around seven million in the human genome [8]. Thus, in GA studies, selecting a subset of SNPs that is sufficiently informative but still small enough to reduce the genotyping overhead is an important step towards disease-gene associations [15]. Machine learning techniques [1] have been applied with ever growing interest and success in GA for feature selection purposes. The success is mostly due to their natural suitability in assessing many loci jointly which may be more informative than testing associations at individual markers and to deal with large amount of data. In fact, most of the studies, in this context, have used a single-locus analysis strategy, in which each variant is tested individually for association with a specific phenotype. However, as often stated in literature, a reason for the lack of success in genetic studies of complex disease is the existence of interactions between loci. If a genetic factor functions primarily through a complex mechanism that involves multiple other genes and, possibly, environmental factors, the effect might be missed if the gene is examined in isolation without allowing for its potential interactions with these other unknown factors [5]. A huge amount of diverse approaches have been developed in the last decade to informative SNPs selection (see [5], [15]) and references therein for an overview) for different disease-gene associations, ranging from multiple sclerosis [3] to chick mortality [11]. Nonetheless, the complexity of GA studies and the peculiarity of each disease-gene association make it impossible to figure out the "best" technique for each application. Random Forest [2], an ensemble of a machine learning techniques based on Decision Trees (DTs), is one candidate model to informative SNP selection. DT [6], is a very popular technique in the machine learning community. DT is efficient from a computational point of view, works well with both categorical or continuous data, copes fine with missing values thanks to the notion of surrogate variables, and its resulting structure after training is easily interpretable even for an user with no machine learning background. On the other hand, DT suffers from high variance. Often a small change in the data can result in a very different tree structure, making interpretation somewhat arbitrary. This motivates the use of RF, which grows a set of many DTs and combines their results through bagging [6] resulting in a very effective reduction of variance in each tree in the ensemble. Aim of this work is a RF informative SNPs selection in Familial Combined Hyperlipidemia (FCH) [12]. FCH is the most common form of familial hyperlipidemia [9]. Affected patients have elevated levels of plasma triglycerides and/or total cholesterol and show increased risk of premature coronary heart disease [13]. In order to identify susceptibility markers for FCH we perform the analysis of 21 SNPs in 10 genes associated with high cardiovascular risk: *Lipoprotein Lipase* (LPL), *Cholesteryl ester transfer protein* (CETP), *HMG-CoA reductase* (HMGCR), PCSK9, ApoA5, ApoC3, *Upstream transcription factor* 1 (USF1), *Peroxisome proliferator-activated recep-*

tor gamma (PPARG), *Gap junction alpha*-4 (GJA4) and *Kinesin-like protein* 6 (KIF6).
The rest of the paper is organized as follows: Section 2 describes the fundamentals of RF, while Section 3 illustrates the FCH sample data used, the experimental settings and the obtained results, respectively. Finally, conclusions in Section 4 close the paper.

2 Random Forests

The RF technique is an ensemble learning technique for conducting classification (or regression) analyses. It constructs a collection of classification trees to aggregate them into one robust classifier. The approach is based on CART (Classification and Regression Trees)[10]. Trees are constructed using rules that determine how well a split at a node (based on the values of a predictor variable) can differentiate observations with respect to the outcome variable. In the following we give an overview of CART principles and motivate the need of an ensemble of CARTs, i.e. the RF in order to better characterize the complexity of our data.

2.1 Classification and Regression Trees

A classification problem consists of a training sample of n observations on a class variable Y assuming values $1, 2, \ldots, k$, and l predictor variables, $X_1, ..., X_l$. The goal is to find a model for predicting the values of Y from new X values. The most classification tree construction algorithms employ a top-down heuristic search using recursive partitioning of data sample space. Starting from a heterogeneous set in terms of the variation in the outcome variable class label (i.e., case or control) of training samples at root node, each predictor (i.e., a SNP) is evaluated using a statistic to determine how well it classifies the training samples by itself. The best feature is selected to split the training samples to descendant nodes. The whole process is recursively repeated to split the descendant nodes until some pre-specified stopping criteria are met (for example, when all terminal nodes contain only cases or only controls, or when all possible SNPs have been included in a branch). The critical step in tree growing is to select the best predictor (or feature in the machine learning jargon) to split a node [4]. Most algorithms evaluate the performance of a candidate feature in separating different class labels in the training samples. The concept of impurity is usually used. Here we consider the Gini index as measure of impurity within node t:

$$I_G(t) = \sum_{i=1}^{k} p_i(t)(1 - p_i(t)),$$

where $p_1, p_2, \ldots, p_k$ are the proportion of samples in the k classes. Then, in binary trees, a feature and a split are chosen according to the following decrement in impurity:

$$\Delta(s, t) = I(t) - h(t_L)I(t_L) - h(t_R)I(t_R)$$

where s is a split of node t, $h(t_L)$ and $h(t_R)$ are the proportions of the samples in the left and right child nodes of node t, respectively. By recursively using the node-splitting procedure, one usually ends up with an overgrown tree (with too many descendant nodes), which produces a tree that overfits the training samples and is prone to random variations in the data. Two commonly employed strategies to overcome the overfitting are to interrupt the tree growing by a stop-splitting criterion and to apply a pruning step on the overgrown tree. The stop-splitting criterion could be either based on the node size, the node homogeneity, or elaborate criterion based on statistical testing.

2.2 Ensemble of trees

Although tree models are easy to interpret, single tree-based analysis has its own limitations in analyzing datasets:

- The topology of a tree is usually unstable. A minor perturbation of the input training sample could result in a totally different tree model. The major reason for this instability is the hierarchical nature of the process: the effect of an error in the top split is propagated down to all of the splits below it.
- For high dimensional data a single parsimonious model is not enough to reflect the complexity in the dataset.
- A single tree may have a relatively lower accuracy in prediction with respect to other machine learning approach (e.g., Support Vector Machine (SVM) or Neural Networks (NN) [1]).

Furthermore, when one faces with small sample sizes, their inherent risk of imprecision and overfitting is higher with single tree models. Rather than using a single classification tree, an high increase in prediction accuracy can result from growing an ensemble of trees and letting them "vote" for the most popular outcome class, given a set of input variable values. This way the ensemble adjusts for the instability of the individual trees induced by small changes in the learning sample. Such ensemble approaches can also be used to provide measures of variable importance. RF is the most widely used ensemble tree approach. A RF consists of hundreds or thousands of unpruned tree each one fitted to a bootstrap sample of the same size (for example, the same number of cases and controls) from the original data. Although an individual tree in the forest is not a good model by itself, the aggregated classification has been shown to achieve much better performance than what a single tree may achieve. To construct an RF with B trees from a training dataset with n observations (cases and controls) with l features or predictors (SNPs), we employ the following steps:

1. A bootstrap sample is drawn from the training sample;
2. A classification tree is grown for the bootstrap sample. At each node, the split is selected on the basis of a randomly selected subset (much smaller than l) features. The tree is grown to full size without pruning;
3. Steps 1 and 2 are repeated B times to form a forest. The ensemble classification label is made by a majority vote of all trees in the ensemble.

Out-f-bag samples, classification accuracy and variable importance The RF algorithm does not use all training samples in the construction of the individual tree, rather it leaves a set of out-of-bag (oob) samples, which can be used to measure the forest classification accuracy [6] with no need of cross-validation or test sets. The out-of-bag observations can also be used to estimate variable importance by a permutation importance index, namely, to measure a specific feature's importance in the tree. When the b-th tree is grown ($b = 1, \ldots B$), the oob samples are passed down the tree, and the prediction accuracy is recorded. The values for the j-th feature are randomly permuted in the oob samples, and the accuracy is again computed. In this way by randomly permuting a feature values, its original association with the outcome variable is broken. When the permuted feature, together with the remaining unpermuted predictor variables, is used to predict the response, the prediction accuracy (i.e., the number of observation classified correctly) decrease substantially if the original variable was associated with the response. Thus, the larger the decrease of the prediction accuracy, the more important the variable. This process is repeated for each feature.

3 Experiments

To detect the genetic loci and their interactions that influence a phenotypic outcome (case or control), the FCH data sample has been analyzed by RF which produces, in each of its constituent tree, a graphical structure (see Fig. 1) that resembles an upside-down tree that maps the possible values of certain predictor variables (e.g., SNP genotypes) to a final expected outcome (e.g., disease status). Each node of the tree represents a predictor variable and there are edges from each node leading down to "child" nodes, in which edge correspond to a different possible value that could be taken by the variable in the parent node. A path through the tree represents a particular combination of values taken by the predictor variables that are present within the path. In this way, we found a subset of SNPs localized in different genes and their association with FCH.
We also performed classical statistical tests for the frequency comparison. Differences between groups were assessed by χ^2 test and a p-value less than 0.05 was considered significant.

3.1 The data

After genomic DNA extraction from peripheral blood samples, the *TaqMan* assay was performed for the SNP typing. Real-Time PCR reactions were performed on a ABI Prism $7900-HT$ instrument and subsequent data analysis were done with the Sequence Detection System 2.3 Software (Applied Biosystem). We enrolled 123 patients with a clinical diagnosis of FCH (after exclusion of misdiagnosis of familial hypercholesterolemia) and 142 healthy controls. Therefore, the case-control data set is constituted by 265 observations and 21 SNPs. The homozygous major, heterozygous, homozygous minor alleles have been coded by 1, 2, and 3,

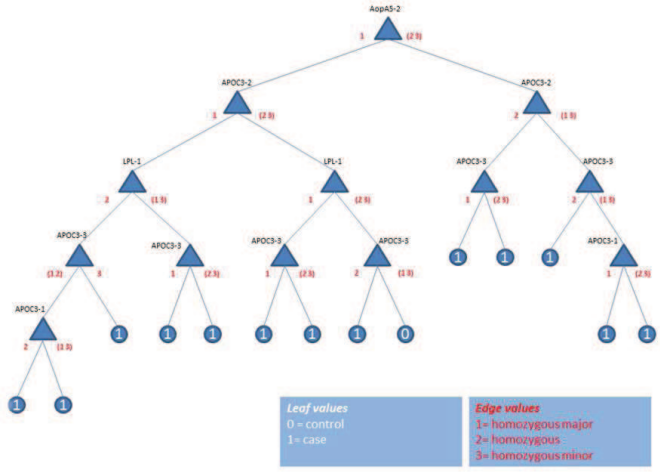

Fig. 1. An example of a single decision tree on case-control data.

respectively. Since our data sample contains few missing values for few SNPs (see Fig.2), the missing genotypes were imputed as the single most likely values on the basis of the genotype frequencies.

3.2 RF analysis

Starting from the original case-control FCH sample, RF formed by 500 to 1000 trees have been trained in order to assess the SNP importance by the permutation importance measure, as Fig. 3 depicts. It is worth noting that, while in Fig. 3 the results for a RF constituted by 1000 trees are reported, a RF with 500 trees would suffice to obtain a similar and stable results. Successively, the RF was used jointly with a sequential backward selection procedure [1] as a wrapper feature selection method [15]. Specifically, the sequential backward selection starts from a subset of eight SNPs (determined by imposing, empirically, a threshold at 0.1 to the permutation importance measure) and sequentially removes a feature from the full candidate set until the removal of further features increase the RF misclassification rate. The RF prediction accuracy as well as the sequential backward selection are computed on the oob samples.

3.3 Results

The identified optimal feature subset contains 5 SNPs and lead to a RF prediction accuracy of 68.68%: *ApoA5 (S19W)*, *ApoA5 (-1131 T> C)*, *PCSK9 (25467958 T> C)*, *GJA4 (P319S)*, *USF1 (9996 G>A)*.
The sensitivity and specificity were 66.67% and 71.83%, respectively. Fig. 4 depicts the the oob error with respect to the number of trees of the RF and clearly

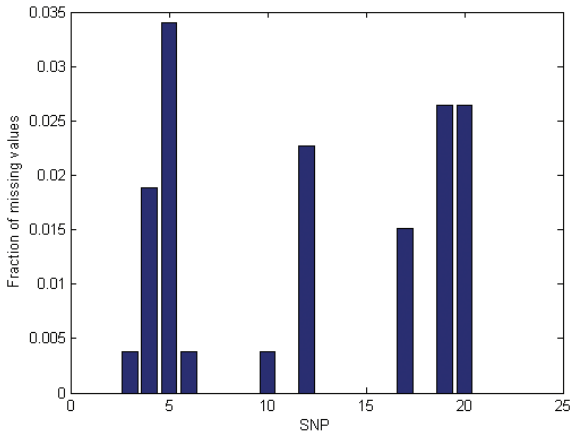

Fig. 2. Fraction of missing values for each SNP in the case-control data set.

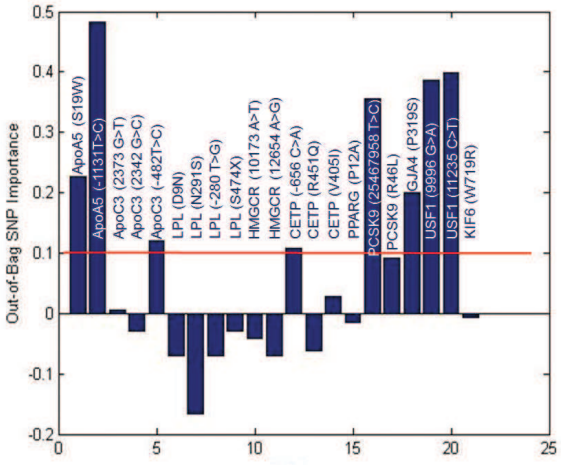

Fig. 3. SNP importance values from permutation importance measure.

shows that a smaller RF size is enough to reach comparable results. Fig. 4 shows the ROC curve obtained by the RF. We also performed standard frequency comparisons to confirm results and we found that genotypes of the above-mentioned SNPs have a different frequency between FCH and controls.

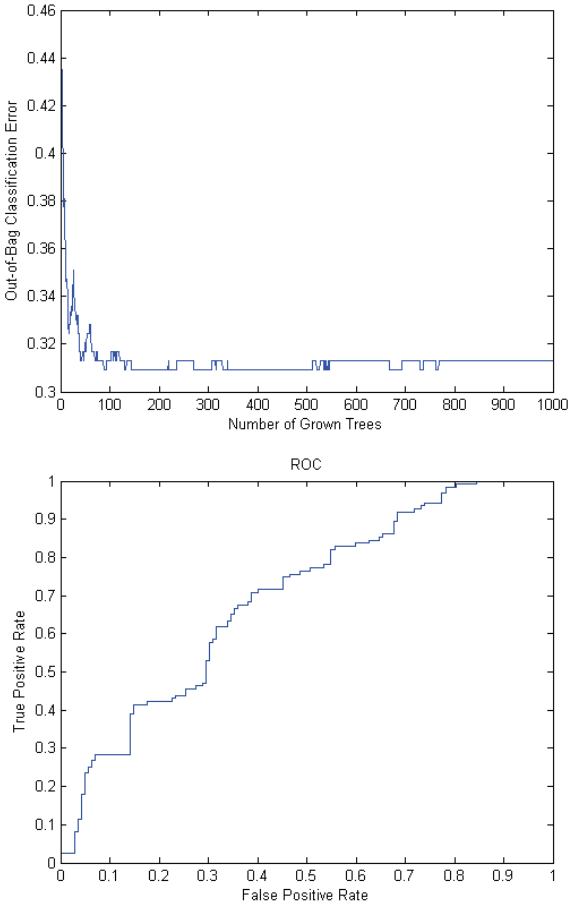

Fig. 4. RF oob error vs number of trees (left) and RF ROC curve (right).

4 Conclusions

RF is an effective methodology for machine learning applications. It is widely used in the research community both for classification and feature selection purposes, due to its some nice properties such as the oob samples which can be

exploited to assess the RF performance with no needs of a testing set or a cross-validation procedure, and to evaluate the feature importance through the permutation importance index. Moreover, an appealing characteristic is the reduced computational costs when compared to other calculus intensive machine learning approach (such as neural networks or support vector machines). Several works in GA literature employ RF as an analysis tool ([14], [16]). We applied a RF analysis to identify a set of SNPs related to Familial Combined Hyperlipidemia: it has been used for a first evaluation of SNP importance through the permutation importance index and then, in conjunction with sequential backward elimination, as a wrapper feature selection procedure that can identify an optimal set on SNPs able to correctly classify 68.68% of subjects in the FCH sample. It is worth stressing that no other employed, for comparison purposes, machine learning approach such as NN and SVM could obtain a prediction accuracy above 65% on the same feature subset. Although the analysis is limited to twenty-one SNPs in ten genes already associated with FCH, the literature reports discordant data about these associations. This motivates our work whose aim was to test the associations with a method of analysis different from the routinely used in case-control studies to confirm the data. So doing, we confirmed the role of the variants $S19W$ and $-1131T > C$ in *ApoA5* gene and $9996G > A$ in *USF1* gene as genetic markers of FCH [9], [12], [13]. Moreover, we reported for the first time association of *GJA4* gene and its variant $P319S$ with FCH. Definitely, while obtained on a relative small sample size, this result could be promising and encourages us to further investigate on the role of this gene in the development of FCH phenotype. Indeed, FCH has polygenic and multifactorial basis not yet completely clarified [13]. Previous linkage and association studies indicated various genes associated to altered lipid or lipoprotein phenotype, most of them being involved in lipid metabolism [7]. Since the polygenic basis of the FCH disease, the combined study of different genetic variations allow to better characterize patients and to identify high-risk subjects. Obviously, the discovery of new variants requires the use of different genetic screening such as wide-genome scan or linkage analysis rather than the small-scale exploratory investigation presented in this work.

Acknowledgements CEINGE Convenzione Regione Campania, DGRC 1901/2009 and IRCCS Fondazione SDN.

References

1. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer (2006)
2. Breiman, L.: Random Forests. Machine Learning. 45(1) (2001)
3. Calcagno, G., Staiano, A., Fortunato, G., Brescia-Morra, V., Salvatore, E., Liguori, R., Capone, S., Filla, A., Longo, G., Sacchetti,L.: A Multilayer Perceptron Neural Network-based Approach for the Identification of Responsiveness to Interferon Therapy in Multiple Sclerosis Patients. Information Sciences. 180(21), 4153–4163 (2010)

4. Chen, X., Wang, M., Zhang, H.: The use of Classification Trees for Bioinformatics. WIREs Data Mining Knowl Discov. 1, 55-63 (2011)
5. Cordell, H.J.: Detecting Genegene Interactions that Underlie Human Diseases. Nat. Revs. Gen. 10, 392–404 (2009)
6. Hastie, J., Tibshirani, R., Friedmanl, J.: The Elements of Statistical Learning - Data Mining, Inference and Prediction. Springer (2009)
7. Hegele, R.A.: Plasma Lipoproteins: Genetic Influences and Clinical Implications. Nat. Revs. Gen. 10(2), 109–121 (2009)
8. Kruglyek, L., Nickerson, D.A.: Variation in the Spice of Life. Nat. Genet. 27, 234–236 (2001)
9. Liu, Z.-K., Hu, M., Baumb, L., Thomas, G.N., Tomlinson, B.: Associations of Polymorphisms in the Apolipoprotein a1/c3/a4/a5 Gene Cluster with Familial Combined Hyperlipidaemia in Hong Kong Chinese. Atherosclerosis 208, 427–432 (2010)
10. Loh, W.-Y.: Classification and Regression Trees. WIREs Data Mining Knowl Discov. 1, 14-23 (2011)
11. Long, N., Gianola, D., Rosa, G.J.M., Weigel, K.A., Avendano, S.: Comparison of Classification Methods for Detecting Associations between SNPs and Chick Mortality. Genet. Sel. Evol. 41(18), (2009)
12. Mar, R., Pajukanta, P., Allayee, H., Groenendijk, M., Dallinga-Thie, G., Krauss, R.M., Sinsheimer, J.S., Cantor, R.M., de Bruin, T.W.A., Lusis, A.J.: Association of the Apolipoprotein a1/c3/a4/a5 Gene Cluster with Triglyceride Levels and Ldl Particle Size in Familial Combined Hyperlipidemia. Circulation Reserch 94(7), 993–999 (2004)
13. Naukkarinen, J., Ehnholm, C., Peltonen, L.: Genetics of Familial Combined Hyperlipidemia. Current Opinion in Lipidology 17, 285–290 (2006)
14. Obulkasim, A., Meijer, G.A., van de Wiel, M.A.: Stepwise Classification of Cancer Samples Using Clinical and Molecular Data. BMC Bioinformatics 12(422), (2011)
15. Saeys, Y., Inza, I., Larranaga, P.: A Review of Feature Selection Techniques in Bioinformatics. Bioinformatics 23(19), 2507–2517 (2007)
16. Yoshida, M., Koike, A.: SNPinterforest: A New Method for Detecting Epistatic Interactions. BMC Bioinformatics 12(469), (2011)

DOI 10.1515/cclm-2014-0144 — Clin Chem Lab Med 2014; x(x): xxx–xxx

Letter to the Editor

Anna Ruotolo, Maria Donata Di Taranto, Maria Nicoletta D'Agostino, Gennaro Marotta, Marco Gentile, Maria Nunziata, Marta Sodano, Rosa Di Noto, Luigi Del Vecchio, Paolo Rubba and Giuliana Fortunato*

The novel variant p.Ser465Leu in the *PCSK9* gene does not account for the decreased LDLR activity in members of a FH family

Keywords: apolipoprotein B (*APOB*); familial hypercholesterolemia; low-density lipoprotein receptor (LDLR) activity; proprotein convertase subtilisin/kexin type 9 (*PCSK9*).

***Corresponding author: Giuliana Fortunato,** Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, via S. Pansini 5, 80131 Napoli; and CEINGE Biotecnologie Avanzate s.c. a r.l., via Gaetano Salvatore 486, 80145 Napoli, Italy, Phone: +39 0817464200, Fax: +39 0817462404, E-mail: fortunat@unina.it
Anna Ruotolo, Maria Nicoletta D'Agostino, Rosa Di Noto and Luigi Del Vecchio: Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, Naples, Italy; and CEINGE Biotecnologie Avanzate s.c. a r.l., Naples, Italy
Maria Donata Di Taranto: IRCCS Fondazione SDN, Naples, Italy
Gennaro Marotta, Marco Gentile, Maria Nunziata, Marta Sodano and Paolo Rubba: Dipartimento di Medicina Clinica e Chirurgia, Università degli Studi di Napoli Federico II, Naples, Italy

To the Editor,

Familial hypercholesterolemia (FH) is a genetically heterogeneous lipid disorder with a frequency of 1:300/1:500 for heterozygotes in many populations [1]. The pathogenesis of FH is caused by a dysfunctional lipid metabolism leading to high concentrations of total and low-density lipoprotein (LDL) cholesterol, however, a great phenotypic variability is observed.

Mutations in the LDL receptor gene (*LDLR*) are the main cause of FH, whereas the apolipoprotein B (*APOB*) and proprotein convertase subtilisin/kexin type 9 (*PCSK9*) genes are involved in a lower percentage of cases [2]. *PCSK9* regulates the *LDLR* expression inducing its lysosomal degradation [3]. Some variants in *PCSK9* increase its capacity to promote LDLR degradation [gain of function mutations (GOF)] leading to FH, whereas other mutations cause a decreased PCSK9 activity [loss of function (LOF)] [3]. The molecular diagnosis of FH implies that a new mutation should be deeply studied and accompanied by a functional characterization before one can claim its pathogenicity. Herein, we present the rare case of a family with a suspect of FH in which two members carried a new variant in the *PCSK9* gene and two members carried a known variant in the *APOB* gene.

The proband (I.2; Figure 1) is a 68-year-old woman with hypercholesterolemia, hypertriglyceridemia as well as a clinical history of hypertension, hyperglycemia, carotid and peripheral atherosclerosis referred to the Dipartimento di Medicina Clinica e Chirurgia, Università degli Studi di Napoli Federico II. The study was performed according to the current version of the Helsinki Declaration. Informed consent was obtained for each patient or control. The proband's father died of a heart attack at the age of 64 and her mother has 7.8 mmol/L of total cholesterol. The patient reported that she suspended previous therapies with atorvastatin or rosuvastatin due to the appearance of myalgia. Laboratory analyses in absence of therapy revealed an altered lipid profile as shown in Table 1. The proband (I.2) showed elevated serum levels of total cholesterol, LDL-cholesterol, triglycerides as well as increased apoB levels and LDL score. No sign of corneal arcus or tendinous xanthomata was found, leading to a diagnosis of possible FH according to the Simon Broome criteria [5]. Sequencing and MLPA analysis of the *LDLR* gene performed as previously described [6] did not reveal any mutation or large rearrangements in the proband. Sequencing of the 12 exons with flanking intron sequences of the *PCSK9* gene showed that the proband (I.2) (Figure 1) was heterozygous for the new variant c.1394C>T in exon 9, corresponding to the amino acid substitution p.Ser465Leu.

We verified the absence of the variant in 150 chromosomes from normocholesterolemic individuals. To evaluate the segregation of the new variant with the

Table 1 Serum lipid profile of the patients.

	I.1[a]	Proband (I.2)[a]	II.1[a]	II.2[a]	Proband (I.2)[b]	II.1[c]	II.2[c]
Total cholesterol, mmol/L	4.2	11.6	6.5	6.2	4.8	5.7	5.9
LDL-cholesterol, mmol/L	2.5	8.4	4.3	4.6	2.3	3.7	4.1
HDL-cholesterol, mmol/L	1.1	1.5	1.4	1.4	1.3	1.5	1.5
Triglycerides, mmol/L	1.2	3.7	1.6	0.6	2.6	1.1	0.5
ApoB, g/L	0.81	1.97	not available	not available	0.9	1.00	1.00
LDL score (sdLDL[d]/LDL)	not available	37.1%	not available	not available	not available	4.2%	3.9%

[a]Values in absence of therapy; [b]Values after 3 years of therapy with simvastatin plus ezetimibe; [c]Values during a low fat diet; [d]sdLDL, small dense LDL; percentage of small dense LDL/total LDL were determined as described in Gentile et al. [4].

hypercholesterolemic phenotype, the mutation was searched in the proband's daughters (II.1 and II.2; Figure 1) aged 43 and 36 who showed elevated total cholesterol and LDL-cholesterol levels only partially corrected by a low fat diet (Table 1) without corneal arcus or tendinous xanthomata. The variant p.Ser465Leu was also identified in one of the proband's daughters (II.2).

To predict the effects of the substitutions we used four different in silico algorithms suggesting that the p.Ser465Leu variant could affect the functionality of the protein: Polyphen (http://genetics.bwh.harvard.edu/pph2) returns the result Probably damaging (score 1.00 at Hum Div and 0.045 at HumVar); SIFT (http://sift.jcvi.org): Not tolerated (score 0.02); Mutation taster (http://www.mutationtaster.org): Disease causing (score 3.95); PMut (http://mmb.pcb.ub.es/): Pathological. BLASTP analysis confirmed that the amino acid residue involved in the substitution is conserved in 15/16 species. These algorithms cannot predict the GOF or LOF effect and it has been reported that predictions of GOF mutations should be interpreted with caution [7]. To assess the effects of the variant, two assays were performed to quantify the residual LDLR activity and the LDLR amount on the surface of patient's cells as an indirect evaluation of the *PCSK9* effect. In fact, PCSK9 is secreted into the plasma and binds directly to LDLR on the cell surface; then, following endocytosis, it triggers LDLR degradation [3].

Peripheral blood mononuclear cells were isolated from patients and controls and incubated for 48 h in a medium with lipoprotein deficient serum (in order to upregulate LDLR) supplemented with ionomycin plus PMA to stimulate T-lymphocytes. Residual activity of LDLR was evaluated by measuring the binding and uptake of a fluorescently labeled LDL (1,19-dioctadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate-conjugated LDL, Life Technology, Foster City, CA, USA), namely DiI-LDL, incubated with the cells for 3 h at 37°C as previously described [8]. Amount of LDLR protein on the cell surface was evaluated by measuring the binding of an anti-LDLR antibody (Progen Biotechnik, Heidelberg, Germany) on stimulated T-lymphocytes. Cells were collected and incubated at 4°C for 1 h with the primary antibody diluted 1:20 followed by an incubation at 4°C for 45 min with the secondary FITC conjugated anti-rabbit antibody (BD Biosciences San Jose, CA, USA) diluted 1:200. Fluorescence intensities were measured with the FACSCanto (Becton-Dickinson,

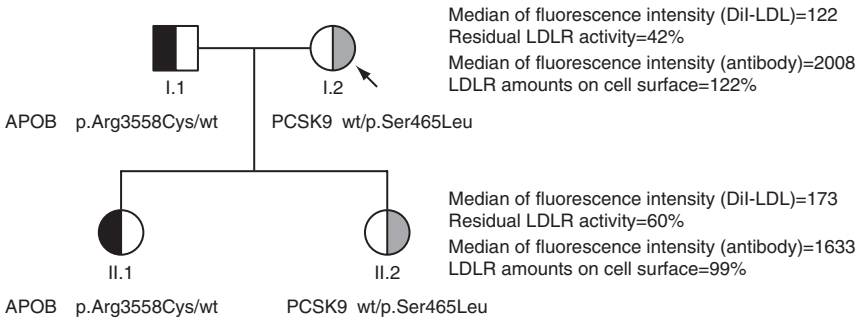

Figure 1 Pedigree of the FH family.
The proband is indicated by an arrow. White and gray symbols indicate heterozygosity for *PCSK9* mutation p.Ser465Leu. Heterozygosity for p.Arg3558Cys mutation in *APOB* is indicated by symbols in black and white. Residual LDLR activities were calculated based on a mean of median fluorescence intensities of the cells from two controls equal to 287; the LDLR amounts were calculated based on a mean of median fluorescence intensities of the cells from two controls obtained with anti-LDLR antibody equal to 1643.

Franklin Lakes, NJ, USA) flow cytometer. LDLR activity or amount of LDLR protein on the cell surface was calculated as the ratio between the median fluorescence intensity of the patients' cells and the mean of median fluorescence intensities of the cells from two controls.

LDLR residual activity in the proband (I.2) and her daughter (II.2) who both bore the c.1394C>T variant, was 42% and 60%, whereas the amounts of LDLR protein on the cell surface were 122% and 99%, respectively (Figure 1).

The decreased LDLR activity observed in both the proband and her daughter (42% and 60%, respectively) together with a small increase or normal LDLR amounts could indicate the dysfunction in LDLR endocytosis as the cause of increased levels of LDL cholesterol in both the mother and the daughter. Our results do not support a pathological role of the new *PCSK9* variant, although it cannot be ruled out because a defect in endocytic pathway could mask its effect, since endocytosis is required for its action as recently described [9]. In addition, other mechanisms of PCSK9 action unrelated to LDR can be responsible for increased cholesterol levels [10].

Furthermore, since the proband's daughter II.1 showed hypercholesterolemia without mutations in *PCSK9*, the genetic screening was extended to the *LDLR* and to all the 29 exons with intron-exon boundaries of *APOB* gene. A mutation in exon 26 of the *APOB* gene (c.10672C>T/p.Arg-3558Cys) was identified in this patient and successively in her father (I.1) who showed a normolipidemic profile.

The p.Arg3558Cys mutation in the *APOB* gene was previously described as pathogenic for FH and associated with a decreased receptor binding affinity of LDL ranging from 30% to 70% and with a high variability of total cholesterol levels ranging from 4.2 mmol/L to 10.9 mmol/L [11]. The above findings are supported by our observations, since p.Arg3558Cys mutation is associated with a normal lipid profile in the patient I.1, but at the same time it is associated with a mild hyperlipidemic profile in the patient II.1.

In conclusion, the new variant p.Ser465Leu in *PCSK9* is not associated to a decreased LDLR amount on the cell membrane and does not explain the reduced LDL activity which is the cause of the FH phenotype. This case report highlights the importance of a deep study of genetic variants before claiming their pathogenicity.

Acknowledgments: We gratefully acknowledge the contribution of the Biobank of CEINGE. Work supported by grants from Regione Campania (Convenzione CEINGE-Regione Campania, DGRC 1901/2009), from IRCCS Fondazione SDN and from POR Campania FSE 2007-2013, Project CREME.

Conflict of interest statement

Authors' conflict of interest disclosure: The authors stated that there are no conflicts of interest regarding the publication of this article. Research support played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.
Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.

Received February 10, 2014; accepted February 12, 2014

References

Q1: References have been renumbered according to the numerical order in text and list. Please check and confirm

1. Hopkins PN, Toth PP, Ballantyne CM, Rader DJ. National lipid association expert panel on familial hypercholesterolemia. Familial hypercholesterolemias: prevalence, genetics, diagnosis and screening recommendations from the National lipid association expert panel on familial hypercholesterolemia. J Clin Lipidol 2011;5:S9–17.
2. Bertolini S, Pisciotta L, Rabacchi C, Cefalù AB, Noto D, Fasano T, et al. Spectrum of mutations and phenotypic expression in patients with autosomal dominant hypercholesterolemia identified in Italy. Atherosclerosis 2013;227:342–8.
3. Tibolla G, Norata GD, Artali R, Meneghetti F, Catapano AL. Proprotein convertase subtilisin/kexin type 9 (PCSK9): from structure function relation to therapeutic inhibition. Nutr Metab Cardiovasc Dis 2011;21:835–43.
4. Gentile M, Panico S, Jossa F, Mattiello A, Ubaldi S, Marotta G, et al. Small dense LDL particles and metabolic syndrome in a sample of middle-aged women. Findings from Progetto Clin Chim Acta 2008;388:179–83.
5. Minhas R, Humphries SE, Qureshi N, Neil HA. Controversies in familial hypercholesterolaemia: recommendations of the NICE guideline development group for the identification and management of familial hypercholesterolaemia. Heart 2009;95:584–7.
6. Romano M, Di Taranto MD, D'Agostino MN, Marotta G, Gentile M, Abate G, et al. Identification and functional characterization of LDLR mutations in familial hypercholesterolemia patients from Southern Italy. Atherosclerosis 2010;210:493–6.
7. Flanagan SE, Patch AM, Ellard S. Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. Genet Test Mol Biomarkers 2010;14:533–7.
8. Romano M, Di Taranto MD, Mirabelli P, D'Agostino MN, Iannuzzi A, Marotta G, et al. An improved method on stimulated

T-lymphocytes to functionally characterize novel and known LDLR mutations. J Lipid Res 2011;52:2095–100.
9. Wang Y, Huang Y, Hobbs HH, Cohen JC. Molecular characterization of proprotein convertase subtilisin/kexin type 9-mediated degradation of the LDLR. J Lipid Res 2012;53:1932–43.
10. Sun XM, Eden ER, Tosi I, Neuwirth CK, Wile D, Naoumova RP, et al. Evidence for effect of mutant PCSK9 on apolipoprotein B secretion as the cause of unusually severe dominant hypercholesterolaemia. Hum Mol Genet 2005;14: 1161–9.
11. Pullinger CR, Gaffney D, Gutierrez MM, Malloy MJ, Schumaker VN, Packard CJ, et al. The apolipoprotein B R3531C mutation: characteristics of 24 subjects from 9 kindreds. J Lipid Res 1999;40:318–27.