UNIVERSITÀ DEGLI STUDI DI PALERMO
Dottorato di Ricerca in Medicina molecolare e Biotecnologie
Curriculum Biotecnologie applicate alla ricerca biomedica
Dipartimento Biomedico di Medicina Interna e Specialistica
SSD MED/09

A NEXT GENERATION SEQUENCING APPROACH FOR MOLECULAR DIAGNOSIS OF MONOGENIC DYSLIPIDEMIAS

IL DOTTORE
DOTT.SSA VALERIA INGRASSIA

IL COORDINATORE
CH.MO PROF. CALOGERO CARUSO

IL TUTOR
CH.MO PROF. MAURIZIO AVERNA

CICLO - XXIX
ANNO DI CONSEGUIMENTO TITOLO - 2017

TABLE OF CONTENTS

INTRODUCTION

Dyslipidemias

Dyslipidemias are heterogeneous group of disorders characterized by abnormal levels of circulating lipids and lipoproteins. These abnormalities may include increases or reductions in cholesterol or elevations in triglycerides. The etiology of these dyslipidemias can further be classified as primary (monogenic) or secondary. Primary or monogenic dyslipidemias are due to single gene defects, while secondary dyslipidemias are multifactorial combining underlying genetic predispositions with disease states such as diabetes, thyroid disease, or drug-related changes in lipid metabolism (1).

Familial hypercholesterolemia: dominant and recessive forms

Monogenic hypercholesterolemias are a group of single gene defects with Mendelian transmission characterized by elevated low-density lipoprotein-cholesterol (LDL-C) levels and very high risk of premature atherosclerotic disease. They are caused by mutations in several genes coding for key proteins involved in endocytosis of the low density lipoprotein receptor (LDLR) and in its pathway (Figure 1) which leads to decrease cellular uptake of LDL with consequent increase in the concentration of LDL-cholesterol (LDL-C) (2).

It is possible to classify the familial hypercholesterolemia in autosomal dominant (ADH) and autosomal recessive (ARH) forms based on the mode of inheritance.

The autosomal dominant hypercholesterolemia (ADH) or classical known as **Familial Hypercholesterolemia (FH)** is one of the most common single gene disorders in the population, is genetically heterogeneous and can be caused by mutations in the following genes:

- the low density lipoprotein receptor gene (LDLR) that cause a reduced bond and catabolism of plasma LDL (autosomal dominant familial hypercholesterolemia type 1, ADH-1). More than 70% of circulating LDL is removed from the blood by LDL-R through hepatic endocytosis. More than 1,300 different mutations of the gene coding

for LDL-R were currently identified in the world (3,4,5), including large rearrangements, single amino acid substitutions, mutations in the promoter region.

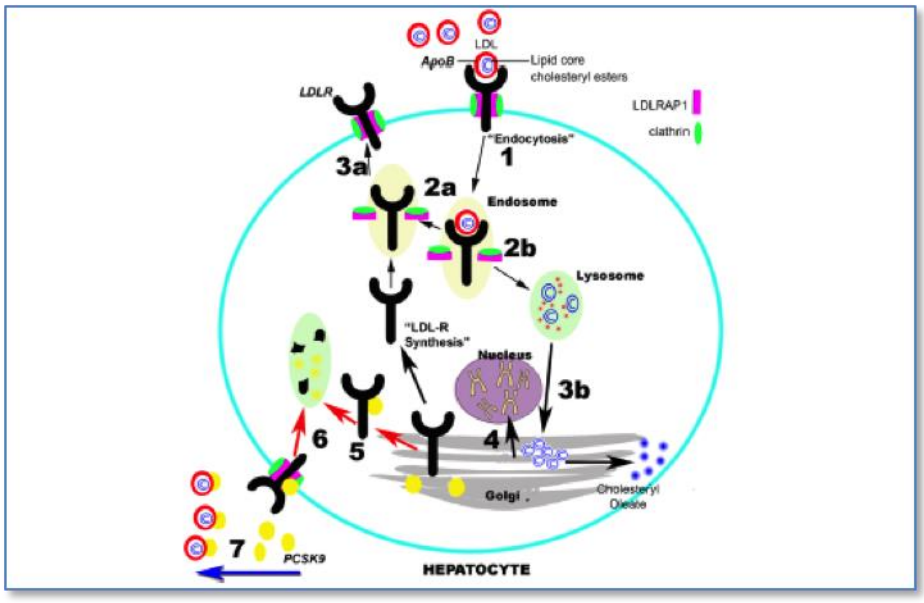

Figure 1. The LDL-R pathway [Adapted from *Henderson et al. Journal of Biomedical Science (2016) 23:39]*. The LDL receptor (LDL-R), part of a LDL-R/clathrin/LDLRAP1 vesicle, binds to the ApoB in LDL particles, internalsing them (1). The receptor-ligand complex dissociates and LDL-R is either recycled (2a and 3a) or degraded (2b and 3b). PCSK9 directs bound LDL-R to the lysosome for degradation (6).

- the apolipoprotein B-100 gene (ApoB) resulting in the production of a defective ApoB-100 which has a reduced binding affinity for LDL-R (Familial hypercholesterolemia Defective ApoB-100, FDB or ADH-2). At the time, four allelic variants of this gene were identified (6). The most common mutation in Europe is Arg3527Gln. 2-5% of patients with hypercholesterolemia are heterozygous for the defective allele. The penetrance of this allele is not 100% so that patients with FDB have less severe phenotype than those with FH from LDLR defect (7, 8).
- the proprotein convertase subtilisin gene / kexin type 9 (PCSK9) with changes in the normal enzyme function PCSK9 proteolytic (Autosomal dominant familial hypercholesterolemia type 3, ADH-3). PCSK9 acts by reducing the amount of LDL-R protein in hepatocytes (9, 10, 11, 12, 13). The first evidence that has provided knowledge on the molecular mechanisms underlying this repression comes from the demonstration that PCSK9 is secreted from the cells and extracellular PCSK9 is then internalized along with LDLR, thus promoting its degradation (14). The region of interaction between PCSK9 and the secreted extracellular domain of the receptor is the

homology repeated domain EGF-A (Epidermal growth factor-like) (15). It was noted that the gain of function mutation (gain-of-function GOF) of the PCSK9 gene increasing the interaction with the LDL-R of about 5-30 times compared to the wild-type PCSK9 (16, 17, 18) resulting in internalization of LDL-R and accumulation of LDL-C.

The autosomal recessive hypercholesterolemia (ARH) has similar clinical characteristics to the dominant forms but with less severity. The ARH can be caused by mutations in the following genes:

- low density lipoprotein receptor adapter protein 1 (LDLRAP1) gene plays an important role in tissue-specific. It is involved in the stabilization of the LDL receptor binding and in intracellular transfer of LDL-R/LDL complex (19).
- ABCG5 (Sterolin 1) and ABCG8 (Sterolin 2) genes due b-sitosterolaemia (20, 21). These genes belong to the family of transmembrane ATP-binding cassette transporters and their mutations lead to elevated plasma concentrations of plant sterols (22).
- the CYP7A1 gene causes the 7α-hydroxylase deficiency. This gene encodes the enzyme cholesterol 7α -hydroxylase that catalyzes the first step of the catabolism of cholesterol and the formation of bile acids. The mutation leads to a reduction of the synthesis of bile acids with accumulation of cholesterol in the liver determining down-regulation of LDLR and resulting in hypercholesterolemia. It was identified in a family (23) which to date remains the only (6).

  Identification of FH is primarily by clinical diagnosis with subsequent confirmation by genetic testing where possible. A family history of premature CHD, a clinical history of premature CHD, physical examination for xanthomas and corneal arcus and elevated plasma LDLC concentration are all used in diagnosis. These characteristics have been used to develop the most widely used clinical criteria to aid diagnosis: the Dutch Lipid Clinic Network (DLCN) criteria (24). It is possible to formulate a diagnosis of FH certain (score >8), probably (score between 6 and 8), possible (score between 3 and 5) or improbably (score betweeen 0 and 2) (25).

Hypobetalipoproteinemia

Hypobetalipoproteinemias (HBL) represent a heterogeneous group of disorders characterized by reduced plasma levels of total cholesterol (TC), low density lipoprotein-cholesterol (LDL-C) and apolipoprotein B (apoB) below the 5th percentile of the distribution in the population (29). In our population the main parameter of TC is 125 mg/l. HBL may be caused by mutations in several known genes or mutations in unidentified genes (primary monogenic HBL) and by several non genetic factors such as diet, drugs, and disease-related conditions (secondary HBL) (30).

In general, the genetic or acquired mechanisms that may reduce LDL-C and apoB plasma levels and cause primary or secondary HBL, respectively, must alter the production, assembly, secretion, or catabolism of apoB-containing lipoproteins (26, 28). Among apoB-containing lipoproteins, low-density lipoprotein (LDL) and very low-density lipoprotein (VLDL) are involved in endogenous triglycerides transport in the blood; chylomicrons (CM) are responsible for dietary lipids transport from intestine into circulation (Figure 2).

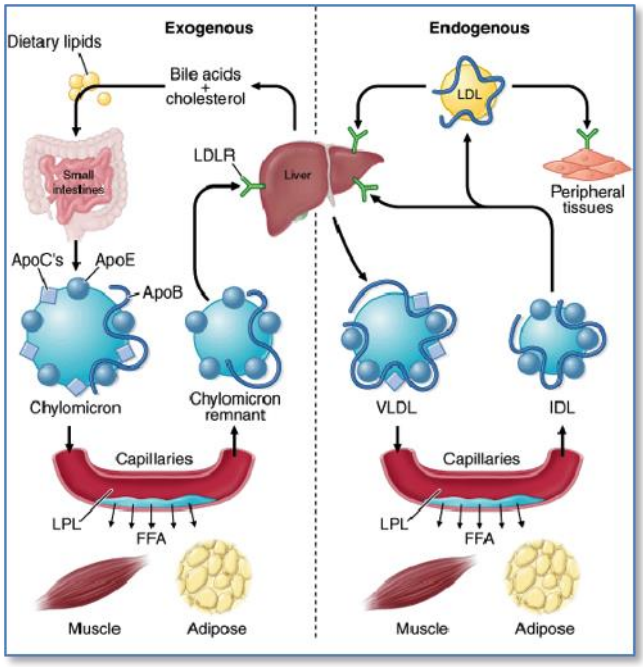

Figure 2. Exogenus and endogenus cholesterol transport pathway. Once inside enterocytes, dietary lipids are packaged into chylomicrons that are exocytosed entering into circulation and distributing fatty acids to muscle and adipose tissues. Endogenous lipids are assembled into VLDL inside hepatocytes and transported to peripheral tissue through circulation.

The apoB protein in plasma is found in two forms, apoB-48 and apoB-100, respectively synthesized in the intestine and liver. The apoB-100 is a structural component of VLDL

rich in triglycerides is the ligand of LDL receptor and is responsible for the clearance of LDL receptor (31). The apoB-48 is required for the "packaging" of the lipids in the chylomicrons.

There are three types of inherited lipoprotein deficiency, characterized by hypocholesterolemia:

a) Familial Hypobetalipoproteinemia (FHBL)
b) Anderson syndrome or "chylomicron Retention Disease" (CRMD);
c) Syndrome of Bassen-Konzweig or Abetalipoproteinemia

Familial hypobetalipoproteinemia (FHBL) is the most frequent monogenic form of HBL with a dominant mode of inheritance and is a genetically heterogeneous disorder. The majority of FHBL patients are heterozygotes while homozygous and compound heterozygous FHBL are very rare (30).

The main candidate gene in subjects with **FHBL** is the APOB gene; approximately 50% of FHBL heterozygotes are carriers of pathogenic mutations in the APOB gene (32). A large number of APOB mutations have been reported to be the cause of FHBL and novel mutations are continually being identified in FHBL subjects (27,28,29). Most APOB gene mutations cause the formation of premature termination codons in the apoB mRNA leading to the formation of truncated apoBs of various size which, to a variable extent, lose the capacity to form plasma lipoproteins in liver and/or intestine and to export lipids from these organs (27,28). Few non-conservative amino acid substitutions in apoB have been reported to be the cause of FHBL (34,35). Two mutations, R463W and L343V, were found to cosegregate with FHBL in two large Libanese kindred. These missense mutations are associated with a decreased secretion of the mutant apoBs that are retained in the ER because of an increased binding to MTP (the chaperone molecule required for apoB lipidation) (36,37). Recently, other missense mutations have been reported to be the cause of FHBL. Five missense APOB mutations located within the N-terminal 1000 amino acids of apoB, namely A31P, G275S, L324M, G912D, and G945S, were identified in heterozygous carriers of FHBL in the Italian population (36).

The large number of FHBL subjects in whom no APOB gene mutations were found suggests that other genes are involved in the pathogenesis of FHBL (28). Another candidate gene for FHBL is PCSK9. Heterozygous subjects carrying the loss-of-function mutations causing PCSK9 truncations were found to have a reduction of plasma LDL-C levels ranging from 30% to 70% (32,33). Loss-of-function mutations of PCSK9 in humans

would increase the receptor mediated uptake and catabolism of plasma LDL, possibly resulting in reduced plasma level of LDL.

Angiopoietin-like protein 3 (ANGPTL3) is a secretory protein regulating plasma lipid levels via affecting lipoprotein lipase and endothelial lipase-mediated hydrolysis of triglycerides and phospholipids. ANGPTL3-deficiency due to loss-of-function mutations in the ANGPTL3 gene causes familial combined hypobetalipoproteinemia **(**also known as **FHBL2)**, a phenotype characterized by low concentration of all major lipoprotein classes in circulation.

The Anderson syndrome or "chylomicron Retention Disease (CMRD) is a rare recessive form of severe hypobetalipoproteinemia charachterized by selective absence of apoB-48. To date 35 cases have been described. Recently, mutations of SARA2 gene encoding the Sar1b protein have been associated with this phenotype. This protein is a regulatory GTP-binding protein involved in intracellular transport of chylomicrons, in particular when Sar1b is activated (Sar1b-GTP) is necessary to complete the polymerization of the coating of COPII vesicles, and its release from the ER. The variants described in the SARA2 gene are missense mutations that cause Sar1b alterations in the conformation of the binding pocket for the GTP, thereby reducing the binding affinity with the same GTP and blocking the formation of COPII vesicle (39).

The other rare recessive form of primary monogenic HBL is represented by **abetalipoproteinemia (ABL)**, the disorder is characterized by extremely low levels of total cholesterol (20-50 mg/dl), LDL-C and apoB. Recent studies have indicated that the molecular defect resides in the gene coding for the "Trygliceride Microsomal Protein Transfer" (MTP). The MTP is expressed at high levels in enterocytes and hepatocytes and its main function is to catalyze the transfer of lipids from their site of synthesis to the apoprotein B (40).

The clinical phenotype of heterozygous FHBL is usually mild, being frequently characterized by fatty liver. The clinical phenotype of homozygous FHBL, ABL, and CMRD is usually severe being characterized by intestinal lipid malabsorption and fat-soluble vitamin deficiency, retinitis pigmentosa, ataxia, neuropathy and acanthocytosis (26,27).

Secondary HBL are due to several non-genetic factors such as diet, drugs, and disease-related conditions. Strict vegetarians (vegans) may have total cholesterol levels approaching the 5th percentile. Intestinal fat malabsorption such as that seen in sprue

(chronic pancreatitis), severe liver disease, malnutrition, and hyperthyroidism may produce low apoB and cholesterol levels (29).

Severe hypertriglyceridemia

Severe hypertriglyceridemia (HTG) is a consequence of a variety of genetic and sporadic metabolic disorders. Many cut offs for severe HTG have been proposed with triglyceride (TG) levels > 1000 mg/dl (11.2 mmol/l) (42) or > 885 mg/dl (10 mmol/l) (41). Mild to moderate HTG is frequently a consequence of common co-morbidities (42) and may have a polygenic etiology with a complex heritability due to the effect of several rare, heterozygous loss-of-function gene variants (43-45). Primary or genetic forms of HTG with a monogenic etiology are much less common and include mostly severe forms characterized by the accumulation in plasma of TG-rich lipoproteins (chylomicrons, VLDL and remnant lipoproteins) (42, 46-47) and an increased risk of developing recurrent episodes of pancreatitis (48). These patients are usually carriers of homozygous or compound heterozygous loss of function mutations of genes pathophysiologically involved in the intravascular lipolysis such as lipoprotein lipase (LPL), apolipoprotein CII (APOC2), apolipoprotein AV (APOA5), glycophosphatidylinositol (GPI)-anchored high-density lipoprotein-binding protein 1 (GPIHBP1), lipase maturation factor 1 (LMF1).

The majority of patients with the monogenic form of HTG (about 51%) is a carrier of the LPL gene mutation, while a small part (16%) is a carrier of rare variants in APOCII, APOA-V and GPIHBP1 (51). LPL has the unique ability to hydrolyse circulating TGs that are packaged in chylomicrons (dietary lipids) or very-low-density lipoproteins (VLDL; lipids of hepatic origin). LPL is present at the cell surface of endothelial cells in small capillaries of tissues requiring fatty acids for their source of energy (heart and skeletal muscle) or for storage (adipose tissue). A complete or near-complete absence of catalytically active LPL invariably leads to severe HTG (51).

In order to have a complete activity, the LPL requires several co-factors (Figure 3): ApoCII and ApoA-V that act as activators, Gpihbp1 (glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1), which binds LPL into the interstitial space and ensures the transportation from the site of synthesis to the capillary lumen through the endothelial cells. The Gpihbp1 also acts as a molecular basis for lipolysis mediated by LPL on the endothelial side of the capillaries. Finally, the active LPL secretion is dependent on the activity of Lmf1, an ER protein consisting of 5 trans-

membrane domains interacting with lipoprotein lipase, promoting its maturation in homodimers before emerging from the ER.

Individuals homozygous or compound heterozygous for mutations such loss-of-function in LPL, ApoCII, ApoA-V, GPIHBP1, or LMF1 show a marked reduction in LPL activity and result in an inappropriate or insufficient clearance of triglycerides in plasma (41-50).

Recently a transient severe form of HTG due to mutations of GPD1 (glycerol-3-phosphate dehydrogenase 1) (42, 48-49) has been described. Although mutations in the LPL gene account for the majority of the cases of primary severe HTG reported so far in literature (49) a growing number of reports describe novel mutations in the other known candidate genes for severe HTG. Moreover, the presence of multiple genetic variants in the major genes affecting LPL-mediated lipolysis of TG-rich lipoproteins can increase the risk of develop severe HTG (44, 49).

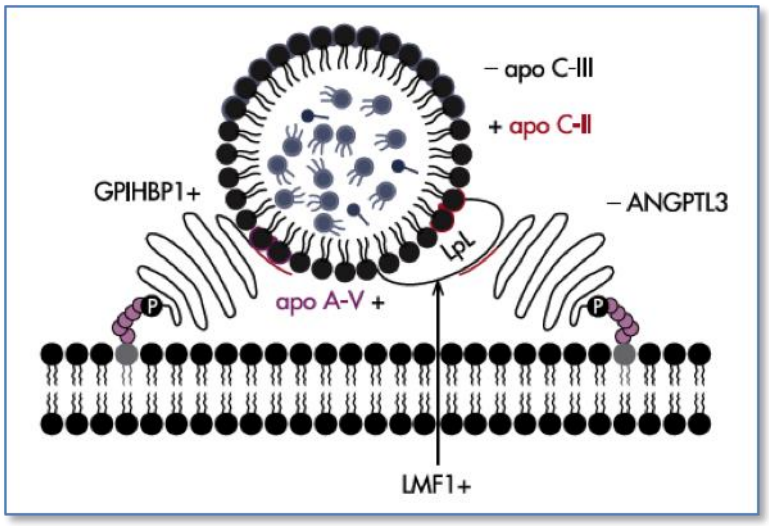

Figure 3. Proteins involved in lipolysis of TRLs in which human mutations result in severe hypertriglyceridemia [Adapted from S. G. Young et al, Curr Opin Lipidol. 2007;18:389–396 (84)]. LMF1 is a chaperone protein for LPL that ensures that LPL attains functional state and is properly secreted. GPIHBP1 tethers LPL to the endothelium, which requires apoC-II as an activator and apoA-V as a stabilizing cofactor. ApoC-III and ANGPTL3 modulates lipolysis near the endothelium.

Dominant familial hypertriglyceridemia recently renamed simple primary hypertriglyceridemia is a very common disorder with an estimated frequency of 1:20 (54) The molecular basis of this form seems to be polygenic (53,55) Recently, the transcription factor cyclic AMPresponsive element–binding protein H (CREB-H, encoded by CRE-binding protein 3–like 3 [CREB3L3]) has been associated with hypertriglyceridemia in humans (52). To date, 4 kindred with dominant hypertriglyceridemia associated with

CREB3L3 heterozygous mutations (245fs and W46X) (52) and one family carrier of the frameshift mutation c.359delG–p.K120fsX20 (47) have been described.

Next generation sequencing

Next generation sequencing (NGS) broadly refers to the new wave of DNA sequencing technologies that have emerged in the post-Sanger sequencing era. NGS platforms provide massively parallel sequencing of millions of DNA fragments, which enables the rapid sequencing of many genes and samples simultaneously (56,58,59). This permits to overcome the limits of classical Sanger method, such as its relative high cost, lengthy time and labor limitations in the clinical diagnostic setting (57). The advantage of these platforms is the determination of the sequence data from amplified single DNA fragments, avoiding the need for cloning of DNA fragments. The major advance offered by NGS is the ability to produce an enormous volume of data cheaply, on the order of hundreds of gigabases of nucleotide sequence per instrument run, while reducing sequencing cost by over five orders of magnitude (60,61,62). Owing to these advantages, NGS technologies have been widely used for many applications and widely adopted in clinical settings. Three main levels of analysis, with increasing degrees of complexity, can now be performed via NGS: disease targeted gene panels, exome sequencing (ES), and genome sequencing (GS) (63). The high-throughput next-generation sequencing (NGS)–based methods have not only proven successful in new disease gene identification, but the availability of economic "benchtop" sequencers offers the realistic possibility to sequence an entire gene or panel of genes, developing molecular diagnoses in inherited disease, especially monogenic diseases, simply and relatively economically (56,57,59).

NGS Platforms

The logistic scheme of the NGS methods can be divided into two main blocks, one related to biology and the second to bioinformatics. In the first, the genome is sequenced following the principle of the *shotgun sequencing*; the genome is amplified fragmented so as to form the so-called *sequencing library*. In the second, the resulting sequence reads are processed through a computational pipeline designed to detect DNA variants.

NGS platforms share a common technological feature massively parallel sequencing of clonally amplified or single DNA molecules that are spatially separated in a flow cell. This design is a paradigm shift from that of Sanger sequencing, which is based on the

electrophoretic separation of chain-termination products produced in individual sequencing reactions. In NGS, sequencing is performed by repeated cycles of polymerase-mediated nucleotide extensions or, in one format, by iterative cycles of oligonucleotide ligation.

As a massively parallel process, NGS generates hundreds of megabases to gigabases of nucleotide sequence output in a single instrument run, depending on the platform (65).

The NGS workflow consists of multiple steps including library preparation, clonal amplification, sequencing and imaging, and data analysis. Extensive reviews have been published on the various NGS technologies (Figure 4), which differ in amplification methods, chemistries, read lengths, throughputs, and run times of different platforms. Commercially available NGS platforms that are suitable for clinical applications include the Roche 454 GS FLX Titanium and Junior systems (Roche Applied Sciences, Penzberg, Germany); and Illumina HiSeq and MiSeq systems (Illumina, San Diego, CA); Life Technology SOLiD, personal genome machine (PGM), and Proton systems (Life Technologies, Carlsbad, CA) (62). Library generation is the process of creating random DNA fragments (100–500 base pairs), of a certain size range, that contain adapter sequences on both ends. The adapters are complementary to platform-specific PCR and sequencing primers. Fragmentation of genomic DNA can be achieved through multiple methods, each having strengths and weaknesses. For most platforms, PCR amplification of the library is necessary before sequencing (63,64).

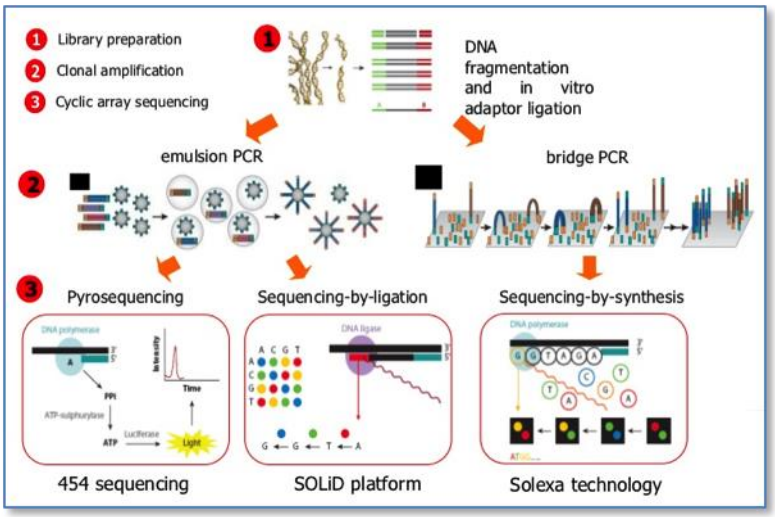

Figure 4. Illustrative scheme of next generation sequencing platforms

The clonal amplification is based on two exclusive methods: emulsion PCR or amplification on solid basis. In emulsion PCR, the single DNA molecules are amplified

clonally in micro-compartments consists of oil and water mixtures. The adapters bound to individual DNA molecules hybridize to complementary sequences coating the surface of specific marbles able to bind covalently to DNA, and to each bead will bind to a single DNA fragment that, by capturing in an emulsion of oil-water, it is amplified in a clonal manner. This amplification method is using by 454, Ion Torrent, and SOLiD technology.

In the Illumina technologies, libraries are denatured and bound at one end to a solid surface coated with adapter oligonucleotides. The free end of each fragment 'bends over' and hybridizes to a complementary adapter on the surface, which initiates complementary strand synthesis. Multiple cycles of this solid-phase amplification followed by denaturation create clusters of ≈1000 copies of single-stranded DNA molecules (64).

The 454 systems (Roche Applied Sciences), work on the principle of pyrosequencing reactions in parallel by loading hundreds of thousands of beads coated with homogeneous DNA fragments into picotiter well plates that are made from fused fiber-optic bundles. The wells are loaded with sequencing enzymes and primer, and then exposed to a flow of one unlabeled nucleotide at a time, allowing synthesis of the complementary DNA strand. When a nucleotide is incorporated, pyrophosphate is released leading to light emission, which is monitored in real time (61).

Illumina (MiSeq or HiSeq) uses sequencing by synthesis chemistry and reversible dye terminator nucleotides. Each sequencing cycle includes incorporation of a nucleotide by DNA polymerase, fluorescence acquisition and nucleotide cleavage (61,64).

SOLiD systems (Life Technologies) use sequencing by ligation technology involving iterative rounds of oligo ligation extension. Each probe is formed by an octamer made of 2 specific bases (in the direction of 3 '→ 5') followed by 6 bases degenerated with one of 4 fluorescent markers bound at the 5' end. In the first step of ligation, a thermostable ligase and the 16 probes representing all possible combinations of 2 bases are used; after binding of the probes, the fluorescence is detected and cleaved the marked portion of the probes so as to regenerate a phosphate group to the 5' and repeat the cycle. This type of approach has the advantage of being able to sequence twice each nucleotide of the template DNA (64).

Ion torrent technology

The Ion Torrent platform (Applied Biosystem) uses the semiconductor sequencing, and it is defined as the third generation technology for further costs reduction produced.

The Ion Torrent technology, in particular the system Ion PGM ™ System - Personal Genome Machine, is a *massive parallel sequencing* technology with exclusive features which differentiate it from other NGS systems.

The PGM ™ sequencer is a "bench top" sequencer, with reduced throughput capacity than competing instruments, with the intention of making cheaper, so more usable, the NGS technology for all projects (targeted DNA sequencing, targeted RNA sequencing, microbial sequencing) where the amount of bases to be sequenced is significantly smaller than the sequencing of an entire human genome.

The Ion Torrent protocol is based on amplicon-based library preparation (Figure 5), which offers the powerful option of sequencing only the regions of interest. For custom panels, primers are designed and typically separated into two pools to minimize unwanted primer-primer interactions. The standard amount of input DNA is 10 ng per pool. After the first round of PCR, sequence-specific primers are removed and the PCR products are phosphorylated. Ion-compatible adapters are then ligated to the amplicons to prepare them for the second round of PCR amplification. The amplicons produced are then purified and quantified for template preparation (61).

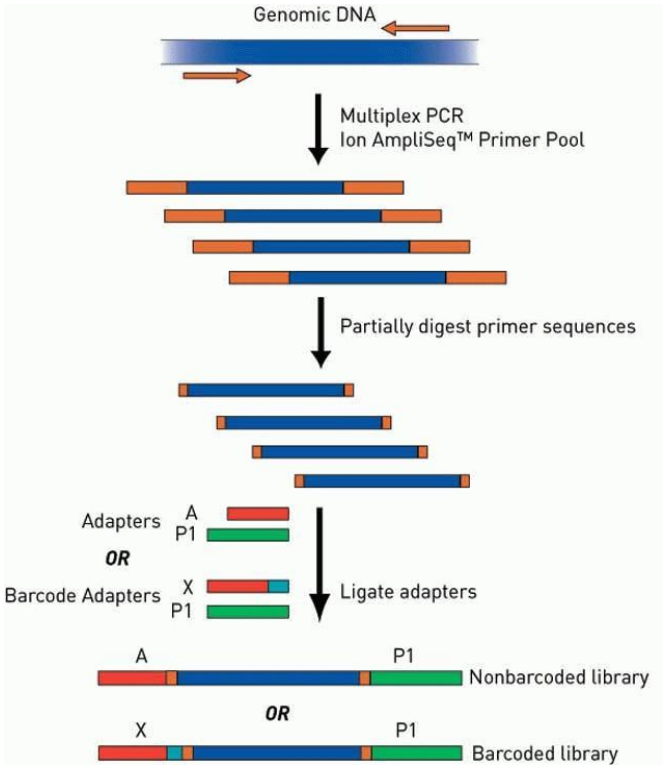

Figure 5. Ampliseq library preparation overview. The genomic DNA is amplified by using specific primer pools in a first PCR, then Ion-specific adapters are added in a second step of PCR producing amplicons ready for the emPCR.

For the PGM sequence platforms the sequence templates are generated on a bead or sphere via emulsion PCR (emPCR). An oil–water emulsion is created to partition small reaction vesicles that each ideally contains one sphere, one library molecule and all the reagents needed for amplification (Figure 6). During the emPCR steps, individual library molecules get amplified to millions of identical copies that are bound to the beads to allow ultimate detection of the signal. In the final step spheres containing amplified DNA are selected in an enrichment step from empty spheres and the loaded spheres are deposited into the sequencing chip (66).

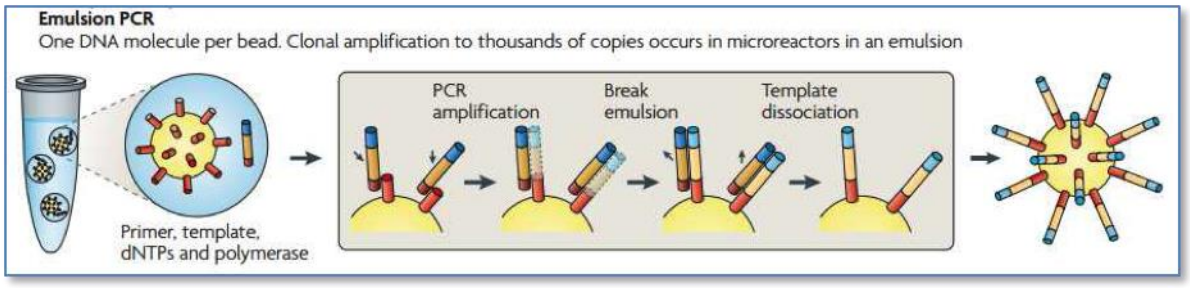

Figure 6. Emulsion PCR workflow (Adapted from M. L. Metzker - Nature Review Genetics 2010). Clonal amplification to thousand of copies occurs in micro-reactors in an emulsion, one oil droplet contains only one DNA molecule.

The Ion torrent chip consists of a flow compartment and solid-state pH sensor micro-arrayed wells that are manufactured using processes built on standard CMOS technology (Figure 7).

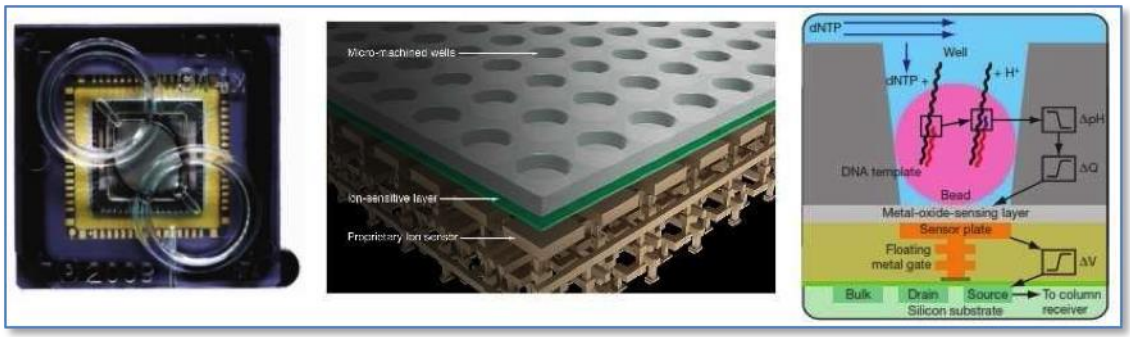

Figure 7. Details of the Ion chip. In the central image is visible a section of the chip with the apparent surface wells. To the right, the signal detection process scheme.

During ion sequencing, all four nucleotides are provided in a stepwise fashion during an automated run. When the nucleotide in the flow is complementary to the template base directly downstream of the sequencing primer, the nucleotide is incorporated into the

nascent strand by the bound polymerase. This increases the length of the sequencing primer by one base (or more, if a homopolymer stretch is directly downstream of the primer) and results in the hydrolysis of the incoming nucleotide triphosphate, which causes the net liberation of a single proton for each nucleotide incorporated during that flow. The release of the proton produces a shift in the pH of the surrounding solution proportional to the number of nucleotides incorporated in the flow (0.02 pH units per single base incorporation) (Figure 8). This is detected by the sensor on the bottom of each well, converted to a voltage and digitized by off-chip electronics (67). This data is represented through a graphic called ionogram, represented in figure 9, which displays the number of bases incorporated respect to flows performed by the device. The advantage offered by this system is the possibility of multiple introductions of bases into the nascent sequence, contrary to what occurs in the Illumina system in which the incorporation for each flow is a single base. However as the number of consecutive identical bases, the Ion Torrent system shows a decrease of accuracy of the bases in the same call (57). The total process of signal generation and measurement (from when the solution containing the nucleotide is released on the surface of the chips until the measurement of pH) lasts little more than 4s. Subsequently, through the flow of a washing solution, necessary to eliminate reagents from the wells already used, the chip is prepared for a subsequent flow that produces the incorporation of the next nucleotide. This process is cyclically repeated for 500 times allowing the reading of long reads about 200bp.

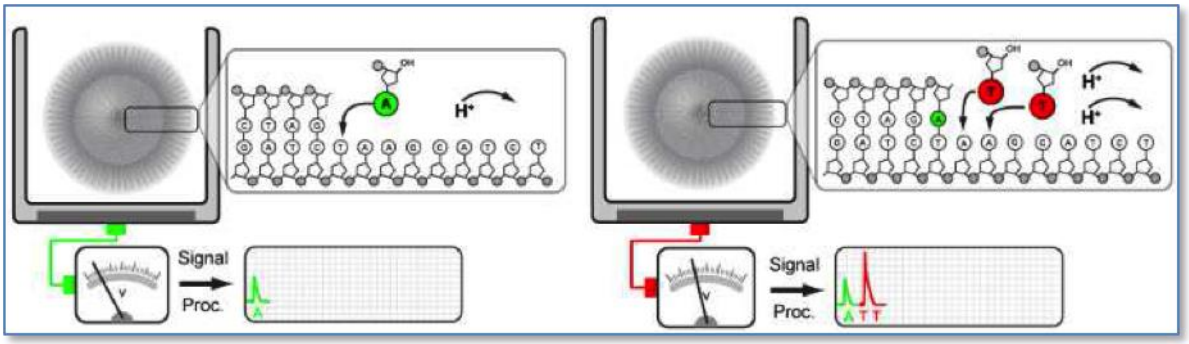

Figure 8. Diagram of the sequencing nell'Ion Chip. Primers, DNA polymerase and beads to which it is linked template are contained inside a well. The four nucleotides cyclic flow on the chip surface. The potential difference is registered for each well to each flow and, if it is registered a change, this is translated into the base call.

Unlike the technologies that use the fluorescence signals, the Ion Torrent system does not provide for the chemical reactions necessary for the elimination of fluorophores groups

linked to nucleotides and this favors the rapidity of the sequencing process that characterizes this technology (68).

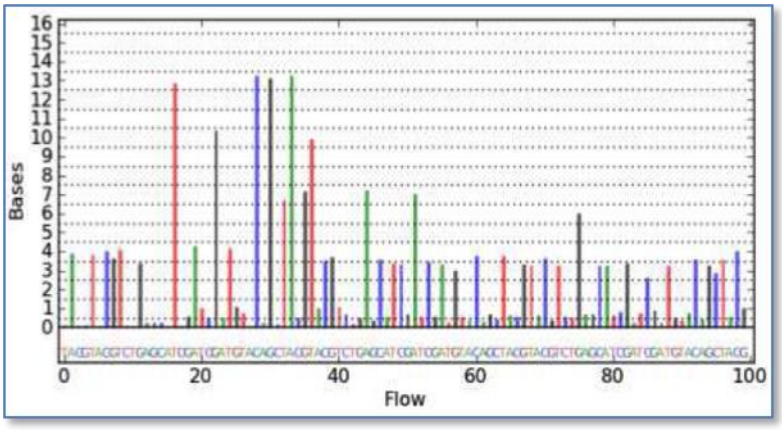

Figure 9. Example of ionogram produced from Ion Torrent PGM™. This graph displays the number of bases incorporated respect to flows performed by the device

An advantageous feature offered by the Ion Torrent technology is the availability of chips with scalar capabilities that enables it to adjust the system according to the experimental needs. In the smaller of the chips currently available for PGM ™ (Ion 314 ™ Chip v2), there are about 1. $2*10^6$ wells, each with a diameter of 3.5 μm, for a total of sequencing capacity ranging between 30 and 100 Mb, while the highest throughput chip (~$12*10^6$ wells) allow sequencing from 600 Mb to 2 Gb (Ion 318 ™ chip v2) (66).

Recently, following the developments implemented to this technology, in addition to "Benchtop PGM™ System", Life Technologies is on the market with the Ion Proton™ Sequencer, an instrument characterized by the same detection technique but capable of higher throughput (10 Gb), more appropriate for massive sequencing of genomes and exomes (66,68).

Bioinformatic analysis

In order to analyze the data, the NGS technology provides a set of computational processes, shown schematically in Figure 10. The multitude of data produced by the NGS platforms is in the order of Terabytes (TB), representing a difficulty for the analysis and storage of data. The analysis software differ according to the NGS technology used in the sequencing, but all follow a system of data analysis pipeline that has the task of converting the luminescence or fluorescence images or current signal in nucleotide sequences ("reads"). In this process, called "base calling", follows the assignment of a rating of

quality ("quality score") to each nucleotide, which indicates the probability of error associated with it. The "quality score" constitute an important tool to eliminate from the analysis process bases or "reads" that do not exceed the appropriate parameters, thereby improving the accuracy of the following "pipeline", such as the alignment of the obtained sequence with reference sequences and variants bases annotation ("variant calls"). For an adequate efficiency of the "pipeline", alignment requires "reads" exceeding 30 bp: in fact, only 90% of the human genome can be aligned uniquely with reads of 30 bp. Another limitation for the alignment is represented by the repeated sequences, whose solution consists in the allocation of sequences repeated "reads" in multiple locations in the reference genome, or in the creation of gaps ("gaps") in the alignment.

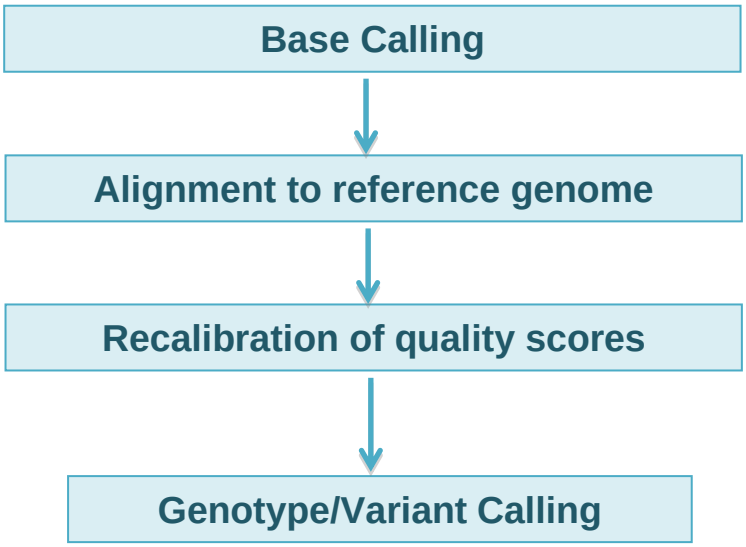

Figure 10. Steps for the analysis generating by NGS

The error rates associated with next-generation technology seem more than the traditional Sanger method; however, the accuracy of sequencing in NGS platforms is ensured by a repeated and massive reading of each gene fragment, which determines the "coverage" of the genome, precisely given by the sum of overlapping "reads" that covering the genome. The latter parameter also represents an essential factor in the analysis, a not suitable "coverage" (in general indicated with value less than 20-50 times in relation to the platform used) can determine the presence of false negatives, for heterozygous samples, in detection of an actual nucleotide variation. At the end of the variants annotation process, the

multitude of variations identified is subject to the application of various filters to reduce the number of candidates.

The most commonly used filters are based on:

- ✓ Inheritance pattern (autosomal / X-linked, dominant / recessive)
- ✓ Equal sharing of variations in well-characterized families
- ✓ Sharing of genes affected by different variations for not related cases
- ✓ Removal or not removal of variants already known through the use of public databases
- ✓ Focusing based on the potential dangers of variant (nonsense, missense, changes in splicing sites or insertions and deletions that alter the reading frame)
- ✓ Prediction of the functional effects of changes through the use of bioinformatics tools *in silico* as SIFT, POLYPHEN, ANNOVAR.

Advantages and limits of NGS technology

The advantages of the technique NGS determine the increase of the data volume produced as a direct result of three factors:

- ✓ many thousands-millions of sequencing reactions can be conducted in parallel overcoming the limit of 1-96 possible reactions with conventional sequencing machines
- ✓ cloning or amplification of the DNA fragment are, in new technologies, unnecessary or fully automated within the platforms
- ✓ capacity to detect with high accuracy the minor allele, which is reflected in the identification of a variation also in mosaic samples or deletions in heterozygosity. In fact the number of times that a DNA fragment is amplified and sequenced is proportional to the abundance of this segment in the original sample, for which the use of specific algorithms in the data analysis can also lead to the identification of changes in the number of copies.

The NGS technology, however, has limitations related mostly to the magnitude of data products; in fact, in the NGS results can be present both false positives and false negatives.

False positives can result from:

- ✓ an incorrect alignment with the reference sequence genomics. It is possible to overcome this limit by the application of data from different alignment software.
- ✓ systematic sequencing errors.
- ✓ errors attributable to the technical limits of the instrument.

False negatives instead come from:

- ✓ the presence of a low coverage
- ✓ low enrichment coverage in regions of interest
- ✓ alignment of repeated regions

The reduction of the number of errors in NGS platforms can be reached through an increase in the coverage and therefore the quality of the run and the fragmentation of the DNA into fragments of greater extension. In fact the short "reads" make it difficult alignment and are difficult to interpret in the determination of the reading step. The Sanger sequencing however shall be required after the analysis of the data, since the results obtained by NGS require validation.

OBJECTIVES

The recent development of applications of next generation sequencing (NGS) has changed the traditional approach of molecular diagnosis based on classical Sanger sequencing methodology, with the possibility to parallelize the sequencing process, which allows the rapid analysis of many genes and samples simultaneously. Thus, the development of next-generation sequencing (NGS) technologies has created new opportunities for the routine use of sequencing in clinical medicine.

The main objective of my work was based on the application of the NGS for the molecular diagnosis of monogenic dyslipidemias characterized by a genetic heterogeneity.

At this purpose, the platform Ion Torrent PGM^{TM} (Thermo Fisher Scientific, Monza, Italia) was applied which performs parallel sequencing by using ion semiconductor chips and the real-time measurements of hydrogen ions produced during DNA replication followed by the direct translation of genetic information (DNA) to digital information (DNA sequence).

The main objective was to evaluate the accuracy and diagnostic yield of three specific targeted NGS panels to establish the molecular diagnosis in a range of patients with high levels of LDL-C, low levels LDL-C and high levels of TG.

METHODS

Samples selection

Targeted sequencing on PGM^{TM} was performed on samples from subjects with familial hypercholesterolemia (TC>200 mg/dl – 5.2 mmol/l), severe hypertriglyceridemia (TG> 885 mg/dl - 10 mmol/l) and hypobetalipoproteinemia (TC<125 mg/dl – 3.23 mmol/l, referred to our Center for clinical and genetic testing.

In total, ninety-seven samples were sequenced by using NGS. In order to validate the technology, six patients, in whom Sanger sequencing had previously identified pathogenic mutations in candidate genes, were included as positive controls.

In detail, forty-nine samples were analyzed from patients with a clinical diagnosis of FH, twenty-nine samples were chosen from patients with a clinical diagnosis of HBL and thirteen were selected from patients clinically diagnosed as having severe HTG. Two positive controls for each group of patients were included in the analysis.

Plasma lipids analysis

Blood samples from probands (10 mL each) were collected into plain and containing EDTA (1 mg/mL) tubes. Plasma TC, TG and HDL-C were measured using standard enzymatic–colorimetric procedures (Instrumentation Laboratory, NY, U.S.A.) in ILAB 300 Plus auto-analyzer Clinical Chemistry System (Instrumentation Laboratory, NY, U.S.A.). LDL-C was calculated by the Friedewald formula only for subjects with TG levels were < 250 mg/dl. LDL cholesterol levels were determined by using Friedewald formula: LDL-C (mg/dl) = TC (mg/dl)- [TG (mg/dl)/5 + HDL-C (mg/dl)].

Table 1. Plasma lipids of hypercholesterolemic subjects selected for NGS sequencing

Subjects FH	Sex M/F	TC (mg/dl)	TG (mg/dl)	HDL (mg/dl)	LDL-C (mg/dl)	Mutations
1 PC	M	302	160	47	223	*LDLR:* c.2390-1G>A (IVS16) Het
2 PC	F	513	147	48	435	*LDLR:* c.1257C>G p.Tyr419X Het *PCSK9:* c.1274A>G p.Asn425Ser Het
3	M	210	126	65	120	*Unknown*
4	M	281	145	56	196	*Unknown*
5	M	420	130	35	359	*Unknown*
6	F	325	61	71	242	*Unknown*
7	F	501	103	79	401	*Unknown*
8	F	227	56	53	137	*Unknown*
9	M	220	111	47	151	*Unknown*
10	F	234	120	52	158	*Unknown*
11	F	225	82	45	164	*Unknown*
12	M	555	213	23	489	*Unknown*
13	F	258	156	45	182	*Unknown*
14	M	415	87	50	348	*Unknown*
15	M	223	148	60	133	*Unknown*
16	M	368	101	60	288	*Unknown*
17	F	261	85	50	194	*Unknown*
18	M	257	102	60	177	*Unknown*
19	F	241	98	52	169	*Unknown*
20	M	458	246	37	372	*Unknown*
21	M	295	503	42	152	*Unknown*
22	F	222	102	86	116	*Unknown*
23	M	216	153	37	148	*Unknown*
24	F	250	50	61	179	*Unknown*
25	F	293	144	54	210	*Unknown*
26	F	264	92	50	196	*Unknown*
27	F	319	90	56	245	*Unknown*
28	F	215	95	36	160	*Unknown*
29	F	226	46	56	161	*Unknown*
30	M	259	126	51	183	*Unknown*
31	F	293	67	53	227	*Unknown*
32	F	300	108	63	215	*Unknown*
33	M	278	169	48	196	*Unknown*
34	F	268	92	61	188	*Unknown*
35	F	278	121	65	189	*Unknown*
36	F	292	157	57	204	*Unknown*
37	M	318	272	31	233	*Unknown*
38	M	236	117	44	169	*Unknown*
39	F	399	63	63	323	*Unknown*
40	F	274	332	34	174	*Unknown*
41	F	255	139	34	193	*Unknown*
42	M	287	160	47	208	*Unknown*
43	F	258	86	84	157	*Unknown*
44	F	278	121	72	182	*Unknown*
45	F	259	199	37	182	*Unknown*
46	F	254	165	27	194	*Unknown*
47	M	250	61	73	165	*Unknown*
48	M	223	87	44	136	*Unknown*
49	M	258	339	25	165	*Unknown*
50	M	278	100	64	194	*Unknown*
51	F	307	118	40	243	*Unknown*

PC, Positive Control; FH, familial hypercholesteromia; M, male; F, female; TC, Total Cholesterol; TG, triglyceride; HDL-C, HDL Cholesterol; LDL-C, LDL cholesterol; Het, heterozygous; Hom, homozygous

Table 2. Plasma lipids of hypocholesterolemic subjects selected for NGS sequencing

Subjects HBL	Sex M/F	TC (mg/dl)	TG (mg/dl)	HDL (mg/dl)	LDL-C (mg/dl)	Mutations
52 PC	F	128	31	90	31,8	*APOB:* c.6718A>T p.Lys2240X Het
53 PC	F	91	50	50	31	*APOB:* c.5993delG p.Ser1929FsX61 Het
54	M	123	60	51	60	*Unknown*
55	M	119	105	34	64	*Unknown*
56	F	119	30	73	40	*Unknown*
57	F	30	21	19	6,8	*Unknown*
58	F	108	42	40	59,6	*Unknown*
59	M	72	162	28	11,6	*Unknown*
60	F	100	83	22	61,4	*Unknown*
61	M	75	125	24	26	*Unknown*
62	M	119	40	45	66	*Unknown*
63	M	120	36	59	53,8	*Unknown*
64	M	107	55	35	61	*Unknown*
65	F	118	40	33	77	*Unknown*
66	M	122	53	45	66,4	*Unknown*
67	M	114	102	37	56,6	*Unknown*
68	F	111	32	35	69,6	*Unknown*
69	M	97	20	59	34	*Unknown*
70	F	119	85	38	64	*Unknown*
71	M	106	35	46	53	*Unknown*
72	F	112	6	72	38,8	*Unknown*
73	M	124	46	46	68,8	*Unknown*
74	F	107	58	43	52,4	*Unknown*
75	M	125	33	61	57,4	*Unknown*
76	M	120	96	41	59,8	*Unknown*
77	F	101	24	55	41,2	*Unknown*
78	M	116	127	39	51,6	*Unknown*
79	M	75	124	34	16,2	*Unknown*
80	M	64	28	56	2,4	*Unknown*
81	M	58	124	5	28,2	*Unknown*
82	M	125	88	43	64,4	*Unknown*

PC, Positive Control; HBL, familial hypobetalipoproteinemia; M, male; F, female; TC, Total Cholesterol; TG, triglyceride; HDL-C, HDL Cholesterol; LDL-C, LDL cholesterol; Het, heterozygous; Hom, homozygous

Table 3. Plasma lipids of hypertriglyceridemic subjects selected for NGS sequencing

Subjects HTG	Sex M/F	TC (mg/dl)	TG (mg/dl)	HDL (mg/dl)	Mutations
83 PC	M	149	920	11	*APOCII:* c.177C>A p.Tyr59X Hom
84 PC	M	482	3800	20	*LPL:* c.686A>G p.His229Arg Het *LPL:* c.829G>A p.Asp277Asn Het
85	F	244	1175	39	*Unknown*
86	M	250	890	28	*Unknown*
87	M	280	1700	32	*Unknown*
88	M	485	1217	27	*Unknown*
89	M	163	2500	26	*Unknown*
90	F	350	1510	25	*Unknown*
91	F	325	1547	45	*Unknown*
92	F	180	860	7	*Unknown*
93	M	250	890	28	*Unknown*
94	M	410	3010	30	*Unknown*
95	M	397	3853	106	*Unknown*
96	M	315	1058	19	*Unknown*
97	F	332	1425	17	*Unknown*

PC, Positive Control; M, male; F, female; TC, Total Cholesterol; TG, triglyceride; HDL-C, HDL Cholesterol. LDL cholesterol is not provided because of TG concentrations were >250 mg/dl and direct LDL cholesterol was no assayed

Design of customized panels for analysis of dyslipidemias related genes

Three different NGS-panels were designed in order to analyze candidate genes of three forms of monogenic dyslipidemias. Primers for each gene were designed using the Web-based Ion AmpliSeq™ Designer software (https://www.ampliseq.com). This tool generates different primer design solutions, which differ for specificity, sample type and application. It was chosen the design pipeline able to give the higher specificity for target genes with an expected coverage of about 97%. Nineteen high-cholesterol-related genes were chosen for affecting different pathways of cholesterol metabolism: LDLR, PCSK9, APOB, SCAP, MTTP, MYLIP, ABCG5, ABCG8, LIPA, FGF19, APOE, LDLRAP1, SREBP1, FGFR4, MIR30C1, MIR30C2, STAP1, CH25H, INSIG2. The final panel consists of 82 kb, targets the coding exons and flanking intron regions of selected genes and include 2 pools, 449 amplicons, with an amplicon length range of 125-275 bp. Thirteen low-cholesterol-related genes were chosen: APOB, PCSK9, MYLIP, ANGPTL3, SORT1, TD26, MTTP, APOC3, ANGPTL4, MIA2, PEMT, PLTP, SAP18. The untraslated region (3' UTR) of two genes (LDLR, LPL) were included in the design as they are considered possible regulatory sites by microRNA. The total genomic region covered consists of a total size of 60 kb, 327 amplicons are obtained with a length range of 125-275 bp.

Eighteen genes related to different pathways of TG metabolism were chosen: LPL, LIPG, LIPC, APOA5, APOC2, APOE, GPIHBP1, LMF1, APOC3, GALNT2, CREB3L3, ANGPTL3, ANGPTL4, ANGPTL8, GCKR, MLXIPL, TRIB1 and GPD1. The final panel

consists of a total size of 62 kb, targets the coding exons and flanking intron regions of selected genes and includes 2 pools, 340 amplicons, with an amplicon length range of 125-275 bp.

Because of high GC content and/or repetitive elements it was impossible to design efficient probes in specific target regions: 8 in FH panel, 8 in HBL panel and 14 in HTG. All of these exons were additionally analyzed by Sanger sequencing as described below.

Table 4. Genes included in the NGS FH-panel

Gene	*Transcript ID*	*Metabolic function*
LDLR	NM_000527	LDL receptor pathway
APOB	NM_000384	LDL receptor pathway
APOE	NM_000041	LDL receptor pathway
LDLRAP1	NM_015627	LDL receptor pathway
PCSK9	NM_174936	LDL receptor pathway
MYLIP	NM_013262	VLDL-LDL secretion
MTTP	NM_000253	VLDL-LDL secretion
LIPA	NM_000235	Intracellular metabolism of cholesteryl esters
ABCG5	NM_022436	Intestinal cholesterol absorption
ABCG8	NM_022437	Intestinal cholesterol absorption
SCAP	NM_012235	Regulation of cholesterol biosynthesis and LDL receptor genes
INSIG2	NM_016133	Regulation of cholesterol biosynthesis and LDL receptor genes
FGF19	NM_005117	Regulation of cholesterol biosynthesis
FGFR4	NM_213647	Regulation of cholesterol biosynthesis
CH25H	NM_003956	Regulation of cholesterol biosynthesis
STAP1	NM_012108	Intracellular signal transduction
SREBP1	NM_001005291	Regulation of cholesterol biosynthesis and LDL receptor genes
MIR30C1	NR_029833	Post-transcriptional regulation of cholesterol related gene expression
MIR30C2	NR_029598	Post-transcriptional regulation of cholesterol related gene expression

Table 5. Genes included in HBL NGS-Panel

Gene	Transcript ID	Metabolic function
APOB	NM_000384	VLDL-LDL secretion
MTTP	NM_000253	VLDL-LDL secretion
PLTP	NM_006227	VLDL-LDL secretion
MIA2	NM_054024	Lipoproteins assembly
PCSK9	NM_174936	LDL receptor degradation
MYLIP	NM_013262	VLDL-LDL secretion
ANGPTL3	NM_014495.3	Inhibitor of lipolysis
ANGPTL4	NM_139314.2	Inhibitor of lipolysis
ANGPTL8	NM_018687.6	Inhibitor of lipolysis
APOC3	NM_000040.1	Inhibitor of lipolysis
SORT1	NM_002959	VLDL-LDL secretion
PEMT	NM_148172	Glycerophospholipid biosynthesis
SAP18	NM_005870	Negative regulation of gene expression
3'UTR LDLR		
3'UTR LPL		

Table 6. Genes included in the NGS HTG-panel

Gene	Transcript ID	Metabolic Function
LPL	NM_000237.2	Intravascular lipolysis
LIPG	NM_006033.2	Intravascular lipolysis
LIPC	NM_000236.2	Intravascular lipolysis
APOA5	NM_001166598.1	Intravascular lipolysis
APOC2	NM_000483	Intravascular lipolysis
APOE	NM_000041.2	Intravascular lipolysis
GPIHBP1	NM_178172.4	Intravascular lipolysis
LMF1	NM_022773.2	Intravascular lipolysis
APOC3	NM_000040.1	ApoCIII plasma level
GALNT2	NM_004481.3	ApoCIII plasma level
CREB3L3	NM_032607.2	Transcription regulation of TG related genes
ANGPTL3	NM_014495.3	Inhibitor of lipolysis
ANGPTL4	NM_139314.2	Inhibitor of lipolysis
ANGPTL8	NM_018687	Inhibitor of lipolysis
GCKR	NM_001486.3	TG and carbohydrate metabolism
MLXIPL	NM_032951.2	TG and carbohydrate metabolism
TRIB1	NM_025195.3	TG and carbohydrate metabolism
GPD1	NM_005276.3	TG and carbohydrate metabolism

Library preparation

Genomic DNA from all subjects were extracted from EDTA treated whole blood samples using the Wizard DNA Purification System (Promega, Italy). DNA quantification was performed using Qubit 2.0 Fluorometer (Invitrogen, Thermo Fisher Scientific, Italy). Targeted regions of FH, HTG and HBL panels were amplified in a Veriti Dx Thermal Cycler (Thermo Fisher Scientific, Italy) by using 10 ng of genomic DNA for each pool, Ion AmpliSeqTM Library Kit 2.0 and 2× Ion AmpliSeq Primers Pool in a total volume of 20 µL. The PCR conditions were 2 min at 99°C for polymerase activation and 20 cycles of 15 sec at 99°C for denaturation and 4 min at 60°C for annealing/extension. The obtained amplicons were treated with FuPa Reagent in a total volume of 22 in order to partially digest the primers and phosphorylate the amplicons. Ion Xpress Barcode adapters were ligated to the amplicons by DNA ligase. The adapters function is to permit the binding on a magnetic bead to make the future sequence, while the barcode is used to distinguish the samples. Barcoded libraries were purified with Agencourt AMPure XP Reagent (Beckman Coulter, CA, USA) and quantified using Ion Library TaqMan Quantitation Kit on the Viia7 Dx Real-Time PCR System (Thermo Fisher).

Clonal amplification

The emulsion PCR was carried out with the Ion One Touch System and Ion PGM Template OT2 200 Kit (Thermo Fisher Scientific, Monza, Italia). Each micro droplet of the emulsion oil/water represents a microreactor, within the amplification occurs. In every single droplet there is a bead, an adapter and (theoretically) a single DNA molecule (monoclonal) or more DNA molecule (polyclonal). The sequencing of the DNA molecule coming from monoclonal spheres will produce sequence data that will be analyzed and processed; all polyclonal spheres instead, will produce sequence data that will not be examined. Then, the template-positive Ion Sphere particles were enriched with Dynabeads MyOne Streptavidin C1 beads on the Ion One touch ES system (Thermo Fisher).

Sequencing and bioinformatics analysis

Sequencing was performed on an Ion PGM™ Sequencer using the Ion PGM 200 Sequencing Kit and the Ion 314 or 316 or 318 Chip depend on the number of loaded samples.

At the end of the run, data were elaborated by Ion Torrent Suite 5.0.2 version and two different plug-ins were applied for a preliminary analysis. Coverage Analysis plug-in (5.0.2.0 version) was used to verify coverage of amplicons, missed bases and a well-adjusted distribution between forward and reverse reads. Variant Calling from the sequencing data was performed with a "variant caller" plug-in (5.0.2.1 version) using the human genome hg19 as reference and germline-low stringency parameter settings.

Only reads with a "mapping quality" >4 and a "base minimum quality score" >10 were included.

Heterozygous variants were called if the minor allele was present in at least 10% of the reads at a given position.

Variants found were annotated in Ion Reporter (version 5.0) from VCFs generated by Ion Torrent through the Annotate Variants Single Sample workflow. The analysis output shows all detected and filtered variants as a list in excel format; for each variant in silico functionality output generated from multiple databases interrogation (SIFT, Poliphen-2, db-SNP, 1000 genomes). In addition ExAC Browser (Beta), Exome Aggregation Consortium (http://exac.broadinstitute.org/) was used; potentially and probably deleterious SNP and INDEL were also visualized by using the Broad Institute Integrative Genomic Viewer (version 2.3.40, www.broadinstitute.org/igv).

Genomic DNA amplification

Genomic DNA regions encompassing suspected pathogenetic mutations in candidate genes analyzed by PGM sequencing were subjected to PCR amplification with suitable primers and conditions.

The amplification profile involves a first step of activation of the polymerase at a temperature of 95°C for 5 min (hot start then the polymerase is inactive at low temperatures). It follows the step of denaturation at 95°C for 1 minute, after which it will have the single-stranded target molecules. At this point the annealing step follows, for 1 minute to different annealing temperatures of the primers, allowing the base-pairing to target specific primers. Finally there is an extension phase at 72°C for 1 minute during

which the polymerase is placed close to the target, recognizes the 3'-OH free of specific primer and begins to polymerize. These steps are repeated for 35 cycles and then switch to a stage of final extension at 72°C for 7 minutes to complete all the amplification products. The PCR reaction was performed in a final volume of 30 μl containing dNTPs (each at a final concentration of 2.5 uM), 3 uL of 10X buffer, 1.5 mM $MgCl_2$, 20 pmol of each primer, 2U of Taq polymerase (Taq Invitrogen) and 100 ng of genomic DNA.
Each amplifier has been evaluated in a standard 1.5% agarose gel using the Sybr Safe, a non-specific DNA intercalating double stranded.

PCR products purification

The fragments obtained by amplification reactions were purified by electrophoresis on agarose gel 1% Low Melting Point in 1X TAE (Tris-acetate 0,04M, 0,001M EDTA) and eluted through purification on column (Wizard® PCR Preps DNA Purification System - Promega). The samples were resuspended in mQ water and quantified by spectrophotometric reading (260 nm); the integrity of the products was evaluated visually on agarose gel 1.5% in TAE.

Sanger sequencing

The DNA sequencing method used is the enzymatic method created by Fred Sanger in the mid-70s, modified over time by the use of Taq DNA polymerase that has favored the development of the cyclic sequencing. The method is based on using a single primer and double-stranded DNA, which is subjected to a series of cycles, each consisting of denaturing, annealing and elongation, in the presence of Taq polymerase, deoxyribonucleotides triphosphate (dNTP) and dideossiribonucleotides triphosphate (ddNTP).
A ddNTP is very similar to its counterpart dNTP, but it is devoid of hydroxy group in the carbon atom in position 3' and the carbon atom in position 2'.
At each cycle DNA fragments will be synthesized which may be interrupted if they have incorporated a ddNTP, not being provided the 3'-OH group for the addition of further nucleotides, and which can be separated on the basis of their length.
The substrate for sequencing DNA is produced by PCR amplification, it converts it into single-stranded form and using it as a template for sequencing.

The four ddNTP are distinguishable among them as labeled with different fluorochromes that emit fluorescence at different wavelength.

The fluorescent DNA fragments are analyzed by an automatic sequencer, which is based capillary electrophoresis, 3500 Genetic Analyzer (Applied Biosystems). The electrophoretic run occurs within 8 capillaries containing a polymer similar to a polyacrylamide gel (POP 7 that allows to separate more or less long fragments). A laser beam hits the capillary exciting the fluorochrome markers fragments of different length. The fluorescence emission is read by a CCD (charge-coupled device) camera. The fluorescence emitted by the excited molecules is collected as a band of a particular wavelength and stored as a digital signal.

To analyze the raw data is used a dedicated program that is able to determine which base corresponds to a given fluorescence intensity (Sequencing Analysis 6) and showing data in a colored graph peaks, each corresponding to a particular base, said electropherogram.

The amplified and purified products, quantified by spectrophotometric reading are used as a template for the direct sequence reaction. A quantity of 10 ng of each template DNA to be sequenced every 100 bp is used, in the presence of 3.2 pM of one of the primers specific for the exon in the analysis (usually the forward primer) and 1 ul of BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystem) that contains the deoxynucleotide, the dideoxynucleotides marked, the reaction buffer and Taq specification in a final volume of 10 ul. The Taq specification has a point mutation in the active site to give good affinity for the ddNTP and a point mutation in the N-terminal domain which eliminates the nuclease activity 5 '→ 3'.

The profile of the reaction sequence involves:

- First step at 96°C for 10 seconds in which the DNA is denatured;
- Second step at 50°C for 5 seconds to allow annealing of the primers ,;
- Third step at 60°C for 4 minutes during which the Taq lengthens the chain until it eventually meets a ddNTP.

This is repeated for 25 cycles and with a rise of temperature of 1 ° C per second. These reactions were purified on plates using BigDye XTerminator® Purification Kit (Applied Biosystems) by adding only two reactants:

- Xterminator ™ Solution eliminates what is not incorporated and free salts after the sequencing reaction;

- SAM ™ Solution improves the performance of the first reagent and also stabilizes the post-purification reactions.

The purified reaction was subjected to capillary electrophoresis for 56 minutes [10 minutes of the capillary with the polymer filler, 10-minute pre-travel to 6-10 mA (constant current) and 36 minutes electrophoresis in 6-10 mA]. The analysis of the raw data was performed using the Sequencing Analysis program 6; the subsequent comparison with the sequences in the database was performed with the program SeqScape 3.

RESULTS AND DISCUSSION

Three customized NGS-panels were generated to capture exons, intron-exons boundaries, and flanking untranslated regions (UTRs) of 19 hypercholesterolemia related genes, 13 hypocholesterolemia responsible genes and 18 genes with known roles on different pathways of TG metabolism, respectively.

Ion PGM sequencing output

The analysis of panels specific genes from selected patients were conducted in different sequencing runs by using the 314 or 316 semiconductor chip.

The sequencing output of the targeted customized genes panel produced in mean a total of 100 Mbases and 600,000 total reads when a chip 314 was used processing a mean of 8 samples (Figure 11). The mean of sequencing output was 653 Mbases and 4,000,000 of total reads when using a chip 316 and analysing a mean of 12 samples simultaneously.

After mapping to the reference human genome (Hg19), a mean of 56,520 (314 chip) or 200,000 (316 chip) mapped reads were obtained for each sample with a mean percentage of reads on target of 98.8 %. Base coverage depth was in the range between 230 and 350 each (uniformity of base coverage 93.32 %) in the run with chip 314 or 650 and 850 after a run performed in chip 316. The mean read length was between 160bp and 180bp (expected range 125-275 bp) and the average amplicon reads by primer pool were equally represented. In the variant calling only the reads with a mapping quality >4 and only bases with minimum quality score >10 were considered.

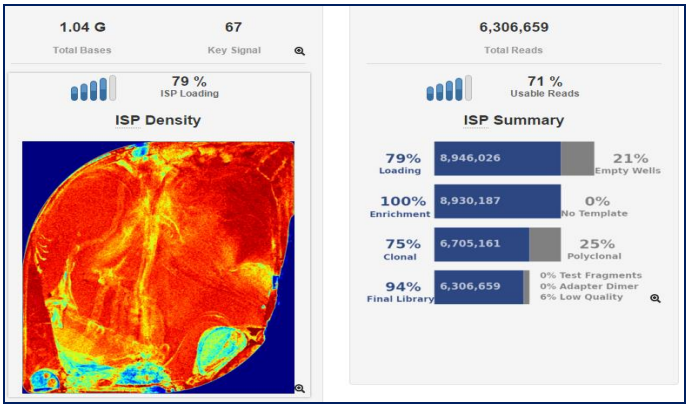

Figure 11. Ion PGM sequencing output using chip 316

Validation of the panels-based NGS

The customized NGS panels were highly accurate and sensitive. The panel pipeline designed with a high specificity for target genes and with coverage of about 97% allowed to confirm genetic variants previously identified by Sanger sequencing in all positive controls. In particular, positive controls were carriers of a variety of mutations in different genes (Table 1, 2, 3) including splicing variants (subject 1), point mutations leading to termination codons (subjects 2, 52, 82), non-synonymous missense (subjects 2, 83) and small in-del leading to reading frame-shift in cDNA and then to the formation of a premature codon stop (subject 53). All the variants previously identified by Sanger sequencing, were redetected giving an analytical sensitivity of 100%.

Detection of variants

The analysis of VCF files obtained after sequencing was performed using the Ion Reporter Software and in particular the plug-in Annotate Variants Single Sample, which allowed generating a list of variants as excel format for each sample.

Forty-nine subjects with a clinical diagnosis of FH were analyzed identifying a mean of 98 variants including 33 exonic variants and 65 non-coding variants. Twenty-nine subjects with a hypocholesterolemic phenotype were studied obtaining for each patient a mean of 65 variants with 25 exonic and 40 intronic variants. In the group of thirteen patients with severe hypertriglyceridemia a mean of 20 exonic and 34 intronic variants were revealed. Among exonic variants synonymous ones were the most common and they were considered as benign and with low priority for Sanger confirmation. Missense non-synonymous variants, nonsense variants, small indels and splicesite variants are respectively listed in Tables 7, 8 and 9 distributed for group of dyslipidemia.

In order to assign a pathogenic role to missense non-synonymous variants the information reported in HGMD (Human Gene Mutation Database) or LOVD (Leiden Open Variation Database) were examined. Variants present within 200 bp of exon/intron boundaries were further evaluated by Human Splicing Finder prediction algorithm in order to compare wild type and variant sequence.

Several identified missense variants were not mentioned in literature and a computer-based analysis (*in-silico* analysis) were performed by using two different algorithms, PolyPhen (www.bork.embl-heidelberg.de/PolyPhen/) and SIFT (http://sift.jcvi.org/), able to predict the effect of amino acid change on protein structure and function.

Analysis of FH patients

Among all patients, one FH (subject 6) resulted carrier of a splicing variant in LDLR receptor (c.313+1G>A IVS3) with a prediction of alteration of the WT donor site by HSF; the mutation is already known and published in literature as pathogenic and responsible of ADH.

Two missense known mutations of APOB gene (c.10579C>T p.Arg3527Trp and c.10708C>T p.His3570Tyr) were identified in subject 9 and 27 respectively described in literature as responsible of FDB. Five missense mutations of LDLR gene already reported in literature as pathogenic and responsible of ADH were identified in 6 subjects which details are described in tables 7.

Four FH patients resulted carriers of unknown missense variants for which the analysis *in silico* indicated a possible deleterious effect for the variants of LDLR and a benign effect for one of PCSK9 (Table 7).

Only one nonsense variants of LDLR gene (c.1257C>G p.Tyr419Ter) was annotated among FH patients, this is already known in literature as ADH disease causing.

One small deletion of 13 bp in LDLR gene (c.316_328delCCCAAGACGTGCT p.Cys109fs) was identified in subject 24, it leads to a formation of a premature stop codon and a truncated non-functional LDLR; this mutation was not found in HGMD or LOVD.

The analysis of the sequence data made also possible to identify an FH proband carrier of an adenine insertion in position 432 of the LDLRAP1 gene. Such insertion causes the formation of a premature stop codon in codon 170, resulting in the formation of a truncated protein. This mutation, known as ARH1, has been identified in homozygosity and is due to a form of autosomal recessive hypercholesterolemia. The visualization of the BAM files through the Integrative Genomics Viewer (IGV) shows ARH1 mutation identified (Figure 12).

To validate the results obtained by NGS, the DNA of the proband was subjected to direct sequencing of exon 4 of LDLRAP1 gene, confirming the presence of the mutation in homozygosity (Figure 13).

The same mutation was previously identified for the first time in a sicilian family carriers (69) and the molecular-genetic characterization of another family carrier of ARH1 mutation together suggest that the autosomal recessive hypercholesterolemia may not be a common disorder only in Sardinia.

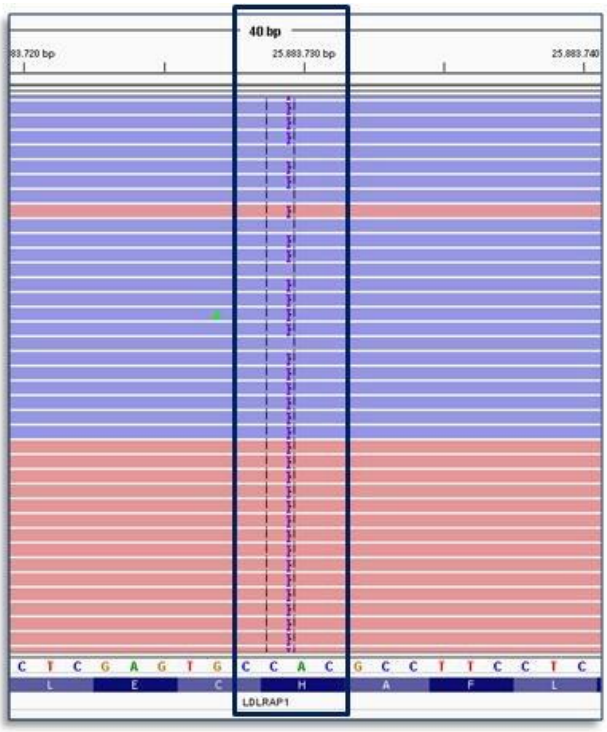

Figure 12. BAM file in IGV visualization of ARH1 mutation c.432insA (p.His144fs) in subject 24

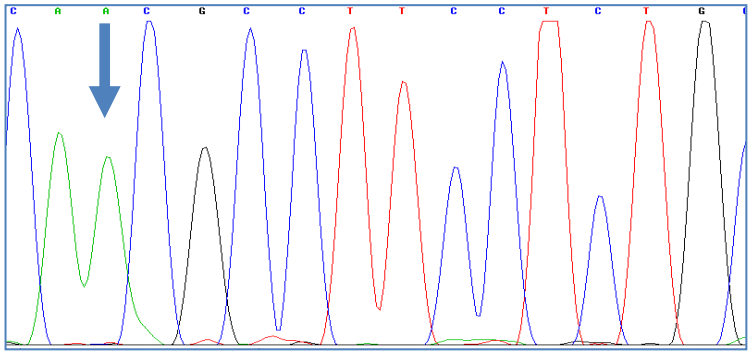

Figure 13. LDLRAP1 c.432insA (p.His144fs)-ARH1 homozygosis

To date there is no data in the literature on ARH1 allele frequency in the population of Sicily; for this reason this represented a proposal to perform a population study for the evaluation of the allele frequency ARH1 in Sicily.

Table 7. Pathogenic mutations identified in FH patients

Gene	Transcript ID	Exon	Mutation	Mutation type	SIFT	POLYPHEN	FH Carrier	Average of reads	Status	MAF	Confirmed by Sanger Sequencing	Ref	Disease or predicted effect
APOB	NM_000384	26	c.10579C>T p.Arg3527Trp	Missense	D	PRB	9	120	Het	0,0004	YES	1	FDB
APOB	NM_000384	26	c.10708C>T p.His3570Tyr	Missense	T	PRB	27	963	Het	0,0002	YES	2	FDB
LDLR	NM_000527	1	c.3G>A p.Met1Ile	Missense	D	POB	31	540	Het	-	YES	-	-
LDLR	NM_000527	2	c.94T>C p.Phe32Ile	Missense	D	POB	29	2436	Het	-	YES	-	-
LDLR	NM_000527	4	c.692G>A p.Cys231Tyr	Missense	T	PRB	50	396	Het	-	YES	-	-
LDLR	NM_000527	6	c.829G>A p.Glu277Lys	Missense	T	B	23	1066	Het	0.0008	YES	3	ADH
LDLR	NM_000527	10	c.1516G>A p.Val506Met	Missense	D	POB	34	176	Het	-		9	ADH
LDLR	NM_000527	10	c.1567G>A p.Val523Met	Missense	D	PRB	16	1103	Het	-	YES	4	ADH
LDLR	NM_000527	11	c.1646G>A p.Gly549Asp	Missense	D	PRB	39, 51	350	Het	0,00005	YES	4	ADH
LDLR	NM_000527	12	c.1729T>C p.Trp577Arg	Missense	D	PRB	4	397	Het	-	YES	5	ADH
PCSK9	NM_174936	9	c.1496G>A p.Arg499His	Missense	T	B	34	140	Het	0,00007	YES	-	-
LDLR	NM_000527	9	c.1257C>G p.Tyr419Ter	Nonsense	-	-	20	1809	Het	-	YES	7	ADH
LDLR	NM_000527	-	c.313+1G/A (IVS3)	3'Splicesite	-	-	6	400	Het	-	YES	8	ADH
LDLR	NM_000527	4	c.316_328delC CCAAGACGT GCT p.C109fs	Small Deletion	-	-	24	1320	Het	-	YES	-	ADH
LDLRAP1	NM_015627	1	c.430_431insA p.His144fs	Small Insertion	-	-	5	203	Hom	-	YES	10	ARH

SIFT, Functional SIFT prediction; T, tolerated; D, damaging. Polyphen, Functional Polyphen Prediction; B, benign; PRB, probably damaging; POB, possibly damaging; MAF, minor allele frequency. MAF is based on 1000 Genomes database. Ref., references. ADH, Autosomal dominant hypercholesterolemia. ARH, Autosomal recessive hypercholesterolemia. FDB, Familial defective APOB100. Het, heterozygous. Hom, Homozygous

Analysis of HBL patients

The evaluation of sequence data in HBL samples permitted to find different kind of small in-del of APOB gene in different patients (Table 8). One small insertion of 6 bp in heterozygosis was identified in exon 26 of APOB gene between positions 10443 and 10444 with a prediction of an in-frame insertion of two aminoacids in the protein chain at position 3481 (subject 55). Since this mutation has not been found described in literature, in vitro functional tests would be needed to test its effect on protein function.

Subject 73 was carrier of a deletion in heterozygosis of 3 bp between the positions 6639 and 6641 of APOB gene, which should remove an aminoacid at position 2213, with uncertain consequence on alteration of apoB protein function needed to further in vitro assays. Further, three different APOB mutations responsible of FHBL were found in three subjects in heterozygous state which prediction on protein is a formation of truncated apoB protein. In particular, the c.10441_10442 insCAAC p.Leu3481fs3525X variant was identified in subject 55 with a predicted truncated apoB of 77,7% in length; the nonsense variant c.10324C>T p.Gln3442Ter identified in subject 56 should result in a protein apoB with an estimated length of 75,8% compared to the wild type one; the subject 80 was carrier of the c.3422delGTCinsTGTGG p.Trp1141fs43X variant which is responsible of the production of an apoB protein of only 1184 aa that is the 26,1% of the wild-type one. As it is known, the most part of FHBL individuals are carriers of APOB gene mutations that cause the formation of premature termination codons in the apoB mRNA leading to the formation of truncated apoBs of various size, which lose the capacity to form plasma lipoproteins in liver and/or intestine and to export lipids from these organs. Thus these three mutations justify the hypocholesterolemic phenotype of respective patients. In particular, in the case of subject 80 it is concern a truncated forms of apoB with a length <30% (apoB-30), which are not generally secreted. The export system of VLDL is altered and results in the accumulation of triglycerides in the liver and it also accomplishes the development of steatosis observed in this subject.

The BAM file in IGV visualization of the mutation c.3422delGTCinsTGTGG p.Trp1141fs43X is reported as example in figure 14 as well as its validation through Sanger sequencing (Figure 15).

All variants were analyzed using Sanger sequencing confirming the presence of them in heterozygous status. Furthermore, they are not reported in HGMD and/or in literature reports.

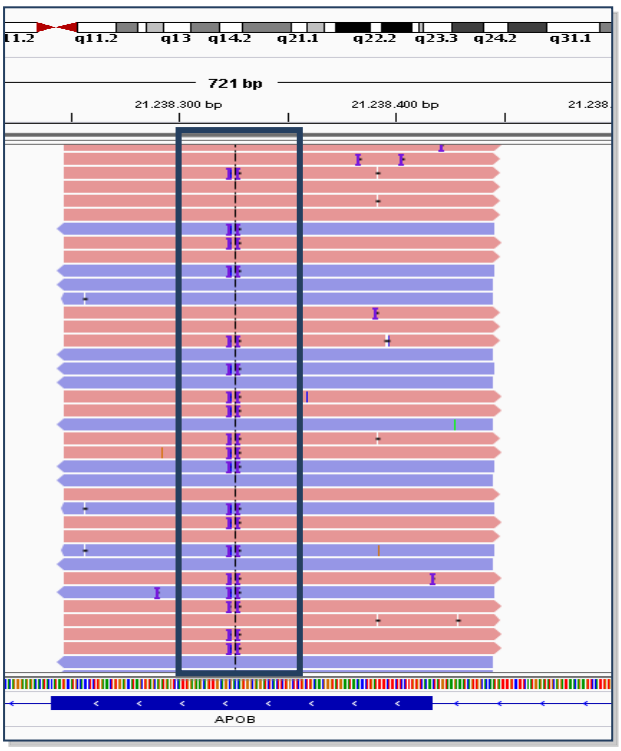

Figure 14. BAM file in IGV visualization of APOB mutation (c.3422delGTCinsTGTGG p.Trp1141fs43X) in subject 80

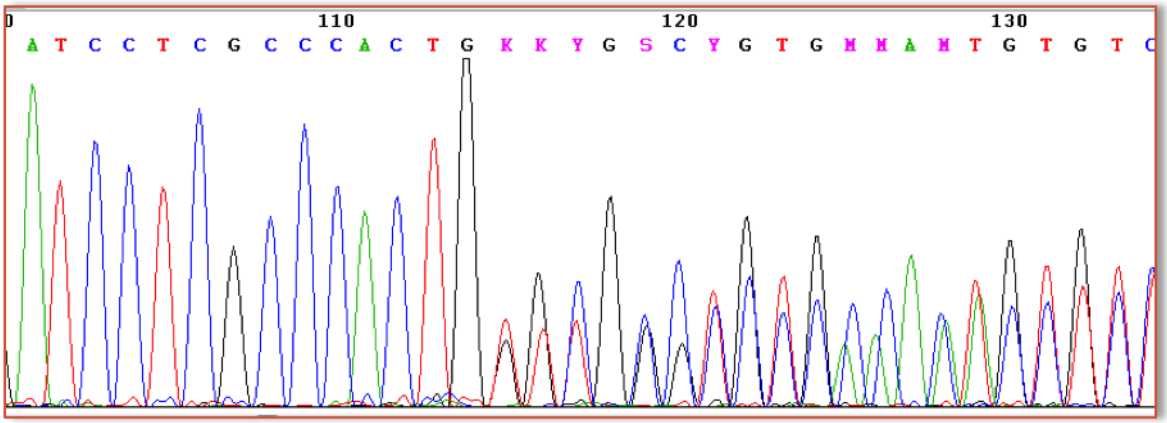

Figure 15. APOB c.3422delGTCinsTGTGG p.Trp1141fs43X heterozygosis

Table 8. Pathogenic mutations identified in HBL patients

Gene	Transcript ID	Exon	Coding	Protein	Mutation type	HBL Carrier	Aver age of reads	Status	Confirmed by Sanger Sequencing	Ref	EFFECT ON PROTEIN	Disease
APOB	NM_000384	26	c.10443_10444 insGTGGAA	p.Leu3481_Ser 3482 insValGlu	Small insertion	55	200	Het	YES	Novel	-	?
APOB	NM_000384	26	c.10441_10442 insCAAC	p.Leu3481fs 3525X	Small insertion	55	180	Het	YES	Novel	Truncated apoB 77,7%	FHBL
APOB	NM_000384	26	c.10324C>T	p.Gln3442Ter	Nonsense	56	120	Het	YES	Novel	Truncated apoB 75,8%	FHBL
APOB	NM_000384		c.6639_6641 delTGA	p.Asp2213del	Small deletion	73	105	Het	YES	Novel	-	?
APOB	NM_000384	22	c.3422delGTC insTGTGG	p.Trp1141fs 43X	Small Deletion/ insertion	80	389	Het	YES	Novel	Truncated apoB 26,1%	FHBL

Analysis of HTG patients

The summary of more relevant annotated variants among HTG patients is listed in table 9.

Subject 86 was found to be carrier of two described missense variants of LMF1 gene (c.1060C>T, p.Arg354Trp and c.1091G>A, p.Arg364Gln). The *in silico* analysis through the two different algorithms, SIFT and Polyphen, indicated a possible deleterious effect on protein function for both variants; however previous in vitro studies did not indicate a significant effect on LPL activity (51).

Two subjects (subject 88 and 92) had a double heterozygous mutation in LPL and in APOA5 genes that may explain the severity of hypertriglyceridemia since other cases of double heterozygosis mutations in HTG candidate genes are described in literature.

Subject 95 was carrier in homozygosis of one missense mutation in LPL gene (c.984G>T p.Met328Ile), this variant was classified as deleterious by SIFT and Polyphen furthermore other patients with severe HTG were described in literature (72).

Subject 97 was a compound heterozygous of two missense variants in heterozygosis in APOA5 gene (c.457G>A p.Val153Met and c.944C>T p.Ala315Val). The first one is classified as polymorphism and the second one does not play any dominant/important role in the genetic determination of plasma TG levels, but the increased frequency in HTG patients compared to controls suggests that it might interact with other gene variants to cause HTG (73).

Subject 93 resulted carrier of a frame-shift mutation in GPIHBP1 gene in homozygosis; it is a deletion of 17 bp at position 413_429 which prevents the formation of stop codon probably leading to a formation of longer protein compared to wild type one thus modifying the action of gpihbp1 protein and than the LPL activity (74).

In subject 89 we identified a homozygous nonsense mutation in LMF1 gene (c.1380C>G, p.Tyr460Ter). Visual inspection of BAM file generated from Ion Torrent PGM sequencing data and visualized with Integrative Genomic Viewer (Figure 16) and Sanger sequencing confirmed the presence of the mutation c.1380C>G in the LMF1 gene (Figure 17). It deals with the identification of the third nonsense mutation of LMF1 gene (c.1380C>G - p.Tyr460ter) in a patient with severe HTG. The c.1380C>G is predicted to generate a premature stop codon and it is expected to be a rare and deleterious variant. It was considering a previous screening analysis of exon 9 of LMF1 gene by direct sequencing performed in 300 unrelated healthy Italian normolipidemic individuals; the mutation was not found in any of these subjects.

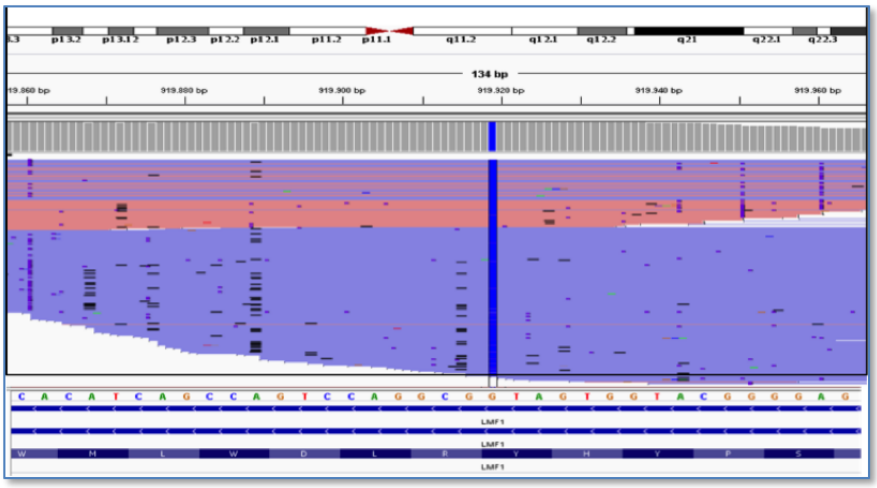

Figure 16. BAM file in IGV visualization of LMF1 mutation (c.1380C>G - p.Tyr460Ter) in subject 89

More, the variant was not reported in the Exome Variant Server repository of the NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS/), in the Exome Aggregation Consortium website (ExAC, http://exac.broadinstitute.org) and the in 1000 Genomes web catalog (http://www.1000genomes.org) and any variant was detected by the analysis of whole LMF1 gene in other 44 unrelated patients with moderate-severe HTG suggesting that c.1380C>G is a rare mutation. The two previously described LMF1 mutations (Tyr439Ter and Trp464Ter) are located in the same domain of Lmf1 protein and both are responsible of a decreased post heparin plasma LPL and HL activity compared with normolipidemic controls (70,71). The nonsense mutation (Tyr460Ter) described here is expected to lead to a truncated Lmf1 protein with an intermediate size compared to the two previously described mutations and a deleterious effect on function can be expected, although any functional studies were performed to confirm this.

Nevertheless, homozygosis for the c.1380C>G mutation resulted in a severe HTG phenotype in the proband.

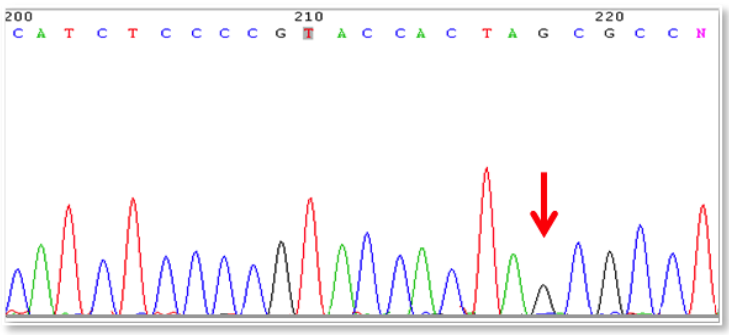

Figure 17. LMF1c.1380C>G - p.Tyr460Ter homozygosis

Table 9. Pathogenic mutations identified in HTG patients

Gene	Transcript ID	Exon	Mutation	Mutation type	SIFT	POLYPHEN	HTG Carrier	Average of reads	Status	MAF	Confirmed by Sanger Sequencing	Ref	Disease
LPL	NM_000237.2	6	c.953A>G p.Asn318Ser	Missense	T	B	88	1542	Het	0,0052	YES	11	HTG
LPL	NM_000237.2	6	c.984G>T p.Met328Ile	Missense	T	B	95	289	Hom	-	YES	12	HTG
LPL	NM_000237.2	8	c.1174C>G p.Leu392Val	Missense	D	PRB	92	130	Het	-	YES	13	Lipoprotein lipase deficiency
APOA5	NM_001166598.1	4	c.427_427delC p.Arg143fs	Small deletion	-	-	88	1053	Het	-	YES	16	HTG
APOA5	NM_001166598.1	4	c.457G>A p.Val153Met	Missense	T	B	92, 97	972	Het Het	0.078	YES	14	Association with HTG
APOA5	NM_001166598.1	4	c.944C>T p.Ala315Val	Missense	D	PRB	97	611	Het	0,004	YES	15	Association with HTG
LMF1	NM_022773.2	7	c.1060C>T p.Arg354Trp	Missense	D	POB	86	767	Het	0.0082	YES	17	No effect
LMF1	NM_022773.2	8	c.1091G>A p.Arg364Gln	Missense	D	POB	86	818	Het	0.0256	YES	17	No effect
LMF1	NM_022773.2	9	c.1380C>G p.Tyr460Ter	Nonsense	-	-	89	1297	Hom	-	YES	-	HTG
GPIHBP1	NM_178172.4	4	c.413_429delTCCCACCCTGGCAAAGC p.Pro140fs	Small deletion	-	-	93	120	Hom	-	YES	18	HTG

SIFT, Functional SIFT prediction; T, tolerated; D, damaging. Polyphen, Functional Polyphen Prediction; B, benign; PRB, probably damaging; POB, possibly damaging; MAF, minor allele frequency. MAF is based on 1000 Genomes database. Ref., references.Het, heterozygous. Hom, Homozygous. HTG, hypertrigliceridemia

Unresolved cases and detection rates

It was considered a molecular diagnosis to have been made if a patient had a pathogenic or probably pathogenic variant detected in a gene known to cause the phenotype present in that patient. Among FH patients, in 15 of a total 49 subjects with a hypercholesterolemia phenotype was possible to define a molecular diagnosis with a detection rate of 30,6 %. Only in 4 patients among 29 with a hypocholesterolemia phenotype it was possible to identify causative mutations in the main HBL candidate gene APOB with a detection rate of 13,8%. The diagnostic yield in the group of HTG subjects with severe HTG was 30,7%, only 4 between 13 patients were carrier of causative mutation in candidate genes.

Even if the diagnostic yield was low, we cannot rule out some technical assay limitation; the designed panels were aimed to detect single nucleotide variants and small indels and by consequence larger indels or complex rearrangements may have been overlooked.

In FH patients who are mutation-negative was recently shown that the clinical phenotype can be associated with an accumulation of common small-effect LDL-C-raising alleles using a 6-SNP score explaining clinical diagnosis of FH with a polygenic cause (75). This should be considered for additional analysis in our negatives subjects.

As regard severe HTG diagnosis, the result confirms previous observations suggesting that the heterogeneity of severe HTG may be explained by a complex interaction between genetic (polygenic burden) and non-genetic factors (43-44).

Evaluation of NGS costs compared with Sanger sequencing

Here, it was described three targeted NGS panels for the molecular diagnosis of disorders characterized by high levels of LDL-C, low levels of LDL-C and high levels of TG. This approach was evaluated analyzing three different groups of patients with a diagnosis of FH, HBL or sever HTG respectively.

Sanger sequencing is the traditional approach used to identify pathogenic mutations in patients with suspected Mendelian disorders of FH, HBL and TG.

The genetic heterogeneity of monogenic dyslipidemias makes this approach costly, labor intensive, time-consuming, and generally requires sequential testing of candidate genes if no pathogenic mutation is identified in the initial screening. In contrast, the targeted re-sequencing strategy, which is described here allows processing of multiple samples and multiple genes in parallel, thus substantially reducing both the costs and processing time.

In table 10 the estimate of costs and processing time of Sanger Sequencing compared to Ion PGM™ Sequencer are shown. The targeted sequencing approach using customized panels appears to be less time consuming and more economical approach for clinical resequencing of FH, HBL and severe HTG.

Table 10. Estimate of costs and processing time of Sanger Sequencing compared to Ion PGM ™ Sequencer

	Sanger Sequencing	Ion PGM ™ Sequencer
Costs of sequencing for 1 patient (€)*	≈ 1300	≈ 280
Estimate of days to acquire, analyze and verify the data ** (1 operator)	90	15

€, Euro; * analysis of total panels genes for 1 patient;
** Time spent to analyze all genes panel for 1 patient

CONCLUSIONS

This study illustrates the use of Next Generation Sequencing in the identification of molecular genetic defects responsible for primary hypercholesterolemia, hypobetalipoproteinemia and hypertriglyceridemia.

Three NGS-panels, specific for each of the three diseases, have been developed for the analysis of candidate genes and they were applied using the Ion Ampliseq technology for the construction of DNA libraries. The Ion One Touch^{TM} system was used for clonal amplification, the semiconductor technology and Ion PGM^{TM} were used for sequencing libraries.

The results showed that customized NGS-panels resulted highly accurate and sensitive for the identification of different type of mutations (missense, nonsense, small indels) allowing molecular diagnosis in several patients with different forms of dyslipidemias.

Moreover, this approach led to detect novel mutations in LDLR and APOB genes and to identification and characterization of the third mutation in the gene LMF1 (c. 1380C> G-p. Tyr460Ter), responsible for the synthesis of an abnormal protein (Lmf1).

This method permitted also to characterize the second Sicilian family affected with autosomal recessive hypercholesterolemia caused by ARH1 mutation.

Furthermore, comparing NGS with the traditional Sanger sequencing approach, it allows processing of multiple samples and multiple genes in parallel, thus substantially reducing both the costs and processing time. In particular, it was estimated a reduction of nearly 50% of costs as well as a reduction in working days increasing the possibility to sequence a higher number of sequenced genes.

In conclusion, the present study highlighted the potential application of Next Generation Sequencing in molecular-genetic diagnostics of different monogenic primitive forms of dyslipidemias.

ACKNOWLEDGMENTS

I would like to thank the Professor Maurizio Averna, for welcoming me in his research group and supervising my research activity during the last three years.

I also want to thank Professor Angelo Baldassare Cefalù for his willingness and professionalism, which supported me with overseeing my work.

Special thanks to my colleagues Rossella Spina and Cinzia Valenti, my point of reference in molecular biology laboratory, for their continue support during my practical work, and for giving me the possibility to establish a relationship of mutual cooperation and trust.

REFERENCES

1. Patni N, Ahmad Z, MD DPW. **Genetics and Dyslipidemia.** In: De Groot LJ, Chrousos G, Dungan K, Feingold KR, Grossman A, Hershman JM, Koch C, Korbonits M, McLachlan R, New M, Purnell J, Rebar R, Singer F, Vinik A, editors. Endotext [Internet]. South Dartmouth (MA): MDText.com, Inc.; 2000-2016 Sep 19.

2. Goldstein JL, Hobbs HH, and Brown MS. **Familial hypercholesterolemia**. in:C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, The Metabolic & Molecular Bases of Inherited Disease. (Eds.), McGraw-Hill, New York, NY, USA, 8th edition, 2001; 2:2863–22913

3. Heath KE, Gahan M, Whittall RA, Humphries SE. **Low-density lipoprotein receptor gene (LDLR) world-wide website in familial hypercholesterolaemia: update, new features and mutation analysis**. Atherosclerosis 2001; 154:243-246

4. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, Abeysinghe S, Krawczak M, Cooper DN. **Human Gene Mutation Database (HGMD): 2003 update**. Hum Mutat 2003; 21:577-581.

5. Usifo E, Leigh SE, Whittall RA, Lench N, Taylor A, Yeats C, Orengo CA, Martin AC, Celli J, Humphries SE. **Low-density lipoprotein receptor gene familial hypercholesterolemia variant database: update and pathologicalassessment.** Ann Hum Genet. 2012; 76:387-401.

6. Soutar AK, Naoumova RP. **Mechanisms of Disease: genetic causes of familial hypercholesterolemia**. Nat Clin Pract Cardiovasc Med 2007; 4:214-225

7. Myant NB. **Familial defective apolipoprotein B-100: a review, including some comparisons with familial hypercholesterolaemia**. Atherosclerosis 1993; 104:1-18.

8. Vrablík M, Ceska R, Horínek A. **Major apolipoprotein B-100 mutations in lipoprotein metabolism and atherosclerosis**. Physiol Res 2001; 50:337-343.

9. Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. **Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9.** Nat Genet 2005; 37:161–165.

10. Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. **Sequence variations in PCSK9, low LDL, and protection against coronary heart disease**. N Engl J Med 2006; 354:1264–1272

11. Hallman DM, Srinivasan SR, Chen W, et al. **Relation of PCSK9 mutations to serum low-density lipoprotein cholesterol in childhood and adulthood (from The Bogalusa Heart Study)**. Am J Cardiol 2007; 100:69–72

12. Hooper AJ, Marais AD, Tanyanyiwa DM, Burnett JR. **The C679X mutation in PCSK9 is present and lowers blood cholesterol in a Southern African population**. Atherosclerosis 2007; 193:445–448

13. Zhao Z, Tuakli-Wosornu Y, Lagace TA, et al. **Molecular characterization of loss-of-function mutations in PCSK9 and identification of a compound heterozygote**. Am J Hum Genet 2006;79:514–523

14. Lagace TA, Curtis DE, Garuti R, et al. **Secreted PCSK9 decreases the number of LDL receptors in hepatocytes and in livers of parabiotic mice**. J Clin Invest 2006;116:2995–3005

15. Li J, Tumanut C, Gavigan JA, et al. **Secreted PCSK9 promotes LDL receptor degradation independently of proteolytic activity**. Biochem J 2007 406:203–207

16. Cunningham D, Danley DE, Geoghegan KF, et al. **Structural and biophysical studies of PCSK9 and its mutants linked to familial hypercholesterolemia**. Nat Struct Mol Biol 2007; 14:413–419

17. Kwon HJ, Lagace TA, McNutt MC, et al. **Molecular basis for LDL receptor recognition by PCSK9**. Proc Natl Acad Sci USA 2008; 105:1820–1825

18. Fisher TS, Lo Surdo P, Pandit S, et al. **Effects of pH and low density lipoprotein (LDL) on PCSK9-dependent LDL receptor regulation**. J Biol Chem2007; 282:20502–20512

19. Garcia CK, Wilund K, Arca M, Zuliani G, Fellin R, Maioli M, Calandra S, BertoliniS, Cossu F, GrishinN, Barnes R, Cohen JC, Hobbs HH. **Autosomal Recessive Hypercholesterolemia Caused by Mutations in a putative LDL Receptor Adaptor Protein**. Science 2001;292:1394-1398

20. Lu K, Lee M-H, Hazard S, Brooks-Wilson A, Hidaka H, Kojima H, Ose L, Stalenhoef AFH, Mietinnen T, Bjorkhem I, Bruckert E, Pandya A, Brewer HB jr, Salen G, Dean M, Srivasta A, Patel SB. **Two genes that map to the STSL locus cause Sitosterolemia: genomic structure and spectrum of mutations involving Sterolin-1 and Sterolin-2, encoded by ABCG5 and ABCG8, respectively**. Am J HumGenet 2001; 69:278-290.

21. Hubacek JA, Beorge KE, Cohen JC, Hobbs HH. **Mutations in ATP-Cassette Binding Protein G5 (ABCG5) and (ABCG8) causing Sitosterolemia**. HumMutat 2001; Mutation in Brief 452.

22. Berge KE, Tian H, Graf GA, Yu L, Grishin NV, Schultz J, Kwiterovich P, Shan B, Barbes R, Hobbs HH. **Accumulation of dietary cholesterol in Sitosterolemia caused by mutations in adjacent ABC transporters**. Science 2000; 290:1771-1775

23. Pullinger CR, Eng C, Salen G, Shefer S, Batta AK, Erickson SK, Verhagen A, Rivera CR, Mulvihill SJ, Malloy MJ, Kane JP. **Human cholesterol 7alpha-hydroxylase (CYP7A1) deficiency has a hypercholesterolemic phenotype**. J ClinInvest2002; 110:109-117

24. Henderson R, O'Kane M, McGilligan V, Watterson S**. The genetics and screening of familial hypercholesterolaemia** J Biomed Sci 2016; 23:39

25. Nordestgaard BG, Chapman MJ, Humphries SE, Ginsberg HN, Masana L, Descamps OS, Wiklund O, Hegele RA, Raal FJ, Defesche JC, Wiegman A, Santos RD, Watts GF, Parhofer KG, Hovingh GK, Kovanen PT, Boileau C, Averna M, Borén J, Bruckert E, Catapano AL, Kuivenhoven JA, Pajukanta P, RayK,Stalenhoef AF, Stroes E, Taskinen MR, Tybjærg-Hansen A; EuropeanAtherosclerosis Society Consensus Panel. **Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease.**

Consensus Statement of the European Atheroclerosis Society. Eur Heart J. 2013; 34:3478-3490

26. Kane JP, Havel RJ. **Disorders of the biogenesis and secretion of lipoproteins containing the B apolipoproteins**, in: C.R. Scriver, A.L. Beaudet, D. Valle, W.S. Sly (Eds.), The Metabolic and Molecular Bases of Inherited Disease, eigth ed., McGraw-Hill, New York, 2001; 2:2717–2752.

27. Hooper AJ, van Bockxmeer FM, Burnett JR, **Monogenic hypocholesterolaemic lipid disorders and apolipoprotein B metabolism**. Crit Rev Clin Lab Sci 2005; 42:515–545.

28. Schonfeld G, Lin X, Yue P. **Familial hypobetalipoproteinemia: genetics and metabolism**. Cell Mol Life Sci 2005; 62:1372–1378.

29. Schonfeld G. **Familial hypobetalipoproteinemia: a review.** J Lipid Res 2003; 44:878-83

30. Tarugi P, Averna M. **Hypobetalipoproteinemia: Genetics, Biochemistry, andClinicalSpectrum.** In Gregory Makowski, editor: Advances in ClinicalChemistry, Burlington: Academic Press, 2011; 54:81-107.

31. Brown MS, Goldstein JL. **Lipoproteinreceptors in the liver. Control signals for plasma cholesteroltraffic.** J Clin Invest 1983; 72:743-747.

32. CohenJ, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. **Low LDL cholesterol in individuals of Africandescentresulting from frequent nonsense mutations in PCSK9**, Nat Genet 2005 37:161–165.

33. Miyake Y, Kimura R, Kokubo Y, et al., **Genetic variants in PCSK9 in the Japanesepopulation: rare genetic variants in PCSK9 might collectively contribute to plasma LDL cholesterol levels in the general population**. Atherosclerosis 2008 196:29–36.

34. Tarugi P, Averna M, Di Leo E, et al. **Molecular diagnosis of hypobetalipoproteinemia: an ENID review**. Atherosclerosis 2007; 195:19–27.

35. Fouchier SW, Sankatsing RR, Peter J, et al. **High frequency of APOB gene mutations causing familial hypobetalipoproteinaemia in patients of Dutch and Spanish descent**. J Med Genet 2005; 42:23.

36. Burnett JR, Shan J, Miskie BA, et al. **A novelnontruncating APOB gene mutation, R463W, causesfamilialhypobetalipoproteinemia**, J Biol Chem 2003; 278:13442–13452.

37. Burnett JR, Zhong S, Jiang ZG, et al. **Missensemutations in APOB within the ba1 domain of human ApoB-100 result in impairedsecretion of ApoB and ApoB-containinglipoproteins in familialhypobetalipoproteinemia**, J Biol Chem 2007; 282:24270–24283.

38. Zhong S, Magnolo AL, Sundaram M, et al. **Non synonymousmutationswithin APOB in human familialhypobetalipoproteinemia: evidence for feedback inhibition of lipogenesis and postendoplasmicreticulumdegradation of apolipoprotein B**, J Biol Chem 2010; 285:6453–6464.

39. Antonny B, Madden D, Hamamoto S, Orci L, Schekman R. **Dynamics of the COPII coat with GTP and stable analogues.** Nat Cell Biol. 2001;3:531-7.

40. Sharp D, Ricci B, Kiennzle B, Lin MC, Wetterau JR. **Human microsomal triglyceride transfer protein large subunit gene structure**. Biochemistry1994; 33: 9057-906

41. Berglund L, Brunzell JD, Goldberg AC, Goldberg IJ, Sacks F, Murad MH, et al. **Evaluation and treatment of hypertriglyceridemia: an Endocrine Society clinical practice guideline.** J Clin Endocrinol Metab 2012; 97:2969-89

42. Hegele RA, Ginsberg HN, Chapman MJ, Nordestgaard BG, Kuivenhoven JA, Averna M, et al. **The polygenic nature of hypertriglyceridaemia: implications for definition, diagnosis, and management.** Lancet Diabetes Endocrinol 2014; 2:655–66

43. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, et al. **Excess of Rare Variants in Non–Genome-Wide Association Study Candidate Genes in Patients With Hypertriglyceridemia.** Circ Cardiovasc Genet 2012; 5:66-72

44. Johansen CT, Hegele RA. **Allelic and phenotypic spectrum of plasma triglycerides.** Biochim Biophys Acta 2012; 1821:833-842

45. Lewis GF, Xiao C, Hegele RA. **Hypertriglyceridemia in the genomic era: a new paradigm.** Endocr Rev 2015; 36:131-47

46. Lee JH, Giannikopoulos P, Duncan SA, Wang J, Johansen CT, Brown JD, et al. **The transcription factor cyclic AMP-responsive element-binding protein H regulates triglyceride metabolism.** NatMed 2011;17:812-5

47. Cefalù AB, Spina R, Noto D, Valenti V, Ingrassia V, Giammanco A, et al. **Novel CREB3L3 Nonsense Mutation in a Family With Dominant Hypertriglyceridemia.** ArteriosclerThrombVascBiol 2015;35:2694-9

48. Brahm AJ, Hegele RA. **Chylomicronaemia-current diagnosis and future therapies.** Nat Rev Endocrinol 2015;11:352-6

49. Rabacchi C, Pisciotta L, Cefalù AB, Noto D, Fresa R, Tarugi P, et al. **Spectrum of mutations of the LPL gene identified in Italy in patients with severe hypertriglyceridemia.** Atherosclerosis 2015:241:79-86

50. Johansen CT, Kathiresan S and Hegele RA. **Genetic Determinants of Plasma Triglycerides**. J Lipid Res 2011; 52:189-206

51. SurendranRP, VisserME, HeemelaarS, Wang J, PeterJ, DefescheJC, KuivenhovenJA, Hosseini M, PeterfyM, KasteleinJJP, JohansenCT,.HegeleR A, StroesESG and Dallinga-ThieGM. **Mutations in LPL, APOC2, APOA5, GPIHBP1 and LMF1 in patients with severe hypertriglyceridaemia.**J Intern Med 2012; 272:185–196

52. Lee JH, Giannikopoulos P, Duncan SA, Wang J, Johansen CT, Brown JD, Plutzky J, Hegele RA, Glimcher LH, Lee AH. **The transcription factor cyclic AMP-responsive element-binding protein H regulates triglyceride metabolism**. Nat Med 2011;17:812–815.

53. Hegele RA, Ginsberg HN, Chapman MJ, et al; **European Atherosclerosis Society Consensus Panel.The polygenic nature of hypertriglyceridaemia: implications for definition, diagnosis, and management.** Lancet Diabetes Endocrinol. 2014;2:655–666.

54. Yuan G, Al-Shali KZ, Hegele RA. **Hypertriglyceridemia: its etiology, effects and treatment.** Can Med Assoc J 2007; 176:1113–1120.

55. Johansen CT, Wang J, McIntyre AD, et al. **Excess of rare variants in nongenome-wide association study candidate genes in patients with hypertriglyceridemia.** Circ Cardiovasc Genet 2012;5:66–72.

56. Johansen CT, Dubé JB, Loyzer MN, MacDonald A, Carter DE, McIntyre AD, Cao H, Wang J, Robinson JF, and Hegele RA. **LipidSeq: a next-generation clinical resequencing panel for monogenic dyslipidemias**. J Lipid Res 2014; 55:765-72

57. Faiz F, Allcock RJ, Hooper AJ, van Bockxmeer FM. **Detection of variations and identifying genomic breakpoints for large deletions in the LDLR by Ion Torrent semiconductor sequencing**. Atherosclerosis 2013; 230:249-255.

58. Grada A and Weinbrecht K. **Next -Generation Sequencing: Methodology and Application.** Journal Invest Dermatol 2013; 133:11

59. Vandrovcova J, Thomas ER, Atanur SS, Norsworthy PJ, Neuwirth C, Tan Y, Kasperaviciute D, Biggs J, Game L, Mueller M, Soutar AK, Aitman TJ. **The use of next-generation sequencing in clinical diagnosis of familial hypercholesterolemia.** Genet Med. 2013; 15:948-957.

60. Ansorge WJ. **Next-generation DNA sequencing techniques.** N Biotechnol 2009; 25: 195-203

61. Chang F, Li MM. C**linical application of amplicon-based next-generation sequencing in cancer.** Cancer Genet 2013; 206:413-419

62. Xuan J, Yu Y, Qing T, Guo L, Shi L. **Next-generation sequencing in the clinic: Promises and challenges.** Cancer Letters 2013; 340:284–295

63. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, Friez MJ, Funke BH, Hegde MR, Lyon E; Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Commitee. **ACMG clinical laboratory standards for next-generation sequencing**. Genet in Med 2013; 15:733-747

64. van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. **Ten years of next-generation sequencing technology.** Trends Genet 2014, 30:418-426

65. Voelkerding KV, Dames SA, Durtschi JD. **Next-Generation Sequencing: From Basic Research to Diagnostics.** Clin Chem 2009; 55:641-658

66. Buermans HP, den Dunnen JT. **Next generation sequencing technology: Advances and applications.** Biochim Biophys Acta 2014;1842-1932

67. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT, Bustillo J. **An integrated semiconductor device enabling non-optical genome sequencing.** Nature 2011; 475:348–352

68. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ. **Performance comparison of benchtop high-throughput sequencing platforms.** Nat biotechnol 2012; 30:434-439

69. Barbagallo CM, Emmanuele G, Cefalù AB, Fiore B, Noto D, Mazzarino MC, Pace A, Brogna A, Rizzo M, Corsini A, Notarbartolo A, Travali S, Averna MR. **Autosomal recessive hypercholesterolemia in a Sicilian kindred harboring the 432insA mutation of the ARH gene**. Atherosclerosis 2003;166:395-400.

70. Cefalù AB, Noto D, Arpi ML, Yin F, Spina R, Hilden H, Barbagallo CM, Carroccio A, Tarugi P, Squatrito S, Vigneri R, Taskinen MR, Péterfy M, Averna MR. **Novel LMF1 nonsense mutation in a patient with severe hypertriglyceridemia**. J Clin Endocrinol Metab 2009; 94:4584-4590

71. Péterfy M, Ben-Zeev O, Mao HZ, Weissglas-Volkov D, Aouizerat BE, Pullinger CR, Frost PH, Kane JP, Malloy MJ, Reue K, Pajukanta P, Doolittle MH. **Mutations in LMF1 cause combined lipase deficiency and severe hypertriglyceridemia.** Nat Genet. 2007; 39:1483-1487

72. Evans D, Arzer J, Aberle J, Beil FU. **Rare variants in the lipoproteinlipase (LPL) gene are common in hypertriglyceridemiabut rare in Type III hyperlipidemia**. Atherosclerosis 2011: 214:386-390

73. Hubacek JA, Wang WW, Skodová Z, Adámková V, Vráblík M, Horínek A, Stulc T, Ceska R, Talmud PJ. **APOA5 Ala315>Val, identified in patients with severe hypertriglyceridemia, is a common mutation with no major effects on plasma lipid levels.** ClinChem Lab Med 46:773-777

74. Gonzaga-Jauregui C, Mir S, Penney S, Jhangiani S, Midgen C, Finegold M, Muzny DM, Wang M, Bacino CA, Gibbs RA, Lupski JR, Kellermayer R, Hanchard NA. **Whole-exome sequencing reveals GPIHBP1 mutations in infantile colitis with severe hypertriglyceridemia.** J Pediatr Gastroenterol Nutr 2014;59:17-21.

75. Futema M, Shah S, Cooper JA, Li K, Whittall RA, Sharifi M, Goldberg O, Drogari E, Mollaki V, Wiegman A, Defesche J, D'Agostino MN, D'Angelo A, Rubba P, Fortunato G, Waluś-Miarka M, Hegele RA, Aderayo Bamimore M, Durst R, Leitersdorf E, Mulder MT, Roeters van Lennep JE, Sijbrands EJ, Whittaker JC, Talmud PJ, Humphries SE. **Refinement of variant selection for the LDL cholesterol genetic risk score in the diagnosis of the polygenic form of clinical familial hypercholesterolemia and replication in samples from 6 countries.** Clin Chem 2015; 61:231-238

TABLES REFERENCES

1. Gaffney D, Reid JM, Cameron IM, Vass K, Caslake MJ, Shepherd J, Packard CJ. **Independent mutations at codon 3500 of the apolipoprotein B gene are associated with hyperlipidemia.** (1995) *Arterioscler Thromb Vasc Biol.* 15:1025-29

2. Soufi M, Sattler AM, Maerz W, Starke A, Herzum M, Maisch B, Schaefer JR. **A new but frequent mutation of apoB-100-apoB His3543Tyr.** *Atherosclerosis*. 2004 174(1):11-6.

3. Pereira E1, Ferreira R, Hermelin B, Thomas G, Bernard C, Bertrand V, Nassiff H, Mendez del Castillo D, Bereziat G, Benlian P. **Recurrent and novel LDL receptor gene mutations causing heterozygous familial hypercholesterolemia in La Habana.** *Hum Genet.* 1995;96(3):319-22.

4. 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.

5. Schmidt HH, Tietge UJ, Buettner J, Barg-Hock H, Offner G, Schweitzer S, Dedoussis GV, Rodeck B, Kallfelz HC, Schlitt HJ, Oldhafer K, Klempnauer J. **Liver transplantation in a subject with familial hypercholesterolemia carrying the homozygous p.W577R LDL-receptor gene mutation.** *Clin Transplant.* 2008; 22(2):180-

6. Bertolini S1, Cassanelli S, Garuti R, Ghisellini M, Simone ML, Rolleri M, Masturzo P, Calandra S. **Analysis of LDL receptor gene mutations in Italian patients with homozygous familial hypercholesterolemia.** Arterioscler Thromb Vasc Biol. 1999;19(2):408-18.

7. Descamps OS1, Gilbeau JP, Leysen X, Van Leuven F, Heller FR. **Impact of genetic defects on atherosclerosis in patients suspected of familial hypercholesterolaemia.** Eur J Clin Invest. 2001 Nov;31(11):958-65.

8. 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 Dec;111(2):175-82.

9. Chmara M, Wasag B, Zuk M, Kubalska J, Wegrzyn A, Bednarska-Makaruk M, Pronicka E, Wehr H, Defesche JC, Rynkiewicz A, Limon J. **Molecular characterization of Polish patients with familial hypercholesterolemia: novel and recurrent LDLR mutations.** J Appl Genet. 2010;51(1):95-106

10. Garcia CK, Wilund K, Arca M, Zuliani G, Fellin R, Maioli M, Calandra S, Bertolini S, Cossu F, Grishin N, Barnes R, Cohen JC, Hobbs HH. **Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein.** Science. 2001 May 18;292(5520):1394-8. Epub 2001 Apr 26.

11. Reymer PW, Gagné E, Groenemeyer BE, Zhang H, Forsyth I, Jansen H, Seidell JC, Kromhout D, Lie KE, Kastelein J, et al. **A lipoprotein lipase mutation (Asn291Ser) is associated with reduced HDL cholesterol levels in premature atherosclerosis.** Nat Genet. 1995 May;10(1):28-34.

12. Evans D, Arzer J, Aberle J, Beil FU. **Rare variants in the lipoprotein lipase (LPL) gene are common in hypertriglyceridemia but rare in Type III hyperlipidemia.** Atherosclerosis. 2011 Feb;214(2):386-90.

13. Pepe G, Chimienti G, Resta F, Di Perna V, Tarricone C, Lovecchio M, Colacicco AM, Capurso A. **A new Italian case of lipoprotein lipase deficiency: a Leu365- > Val change resulting in loss of enzyme activity.** Biochem Biophys Res Commun. 1994 Mar 15;199(2):570-6.

14. Yin RX, Li YY, Liu WY, Zhang L, Wu JZ. **Interactions of the apolipoprotein A5 gene polymorphisms and alcohol consumption on serum lipid levels.** PLoS One. 2011 Mar 14;6(3):e17954.

15. Hubacek JA, Wang WW, Skodová Z, Adámková V, Vráblík M, Horínek A, Stulc T, Ceska R, Talmud PJ. **APOA5 Ala315>Val, identified in patients with severe hypertriglyceridemia, is a common mutation with no major effects on plasma lipid levels.** Clin Chem Lab Med. 2008;46(6):773-7.

16. Evans D, Aberle J, Beil FU. **Resequencing the apolipoprotein A5 (APOA5) gene in patients with various forms of hypertriglyceridemia.** Atherosclerosis. 2011 Dec;219(2):715-20.

17. Surendran RP, Visser ME, Heemelaar S, Wang J, Peter J, Defesche JC, et al. **Mutations in LPL, APOC2, APOA5, GPIHBP1 and LMF1 in patients with severe hypertriglyceridaemia.** J Intern Med 2012;272:185-96

18. Gonzaga-Jauregui C, Mir S, Penney S, Jhangiani S, Midgen C, Finegold M, Muzny DM, Wang M, Bacino CA, Gibbs RA, Lupski JR, Kellermayer R, Hanchard NA. **Whole-exome sequencing reveals GPIHBP1 mutations in infantile colitis with severe hypertriglyceridemia.** J Pediatr Gastroenterol Nutr. 2014 Jul;59(1):17-21