



UNIVERSITÀ DI PARMA

DOTTORATO DI RICERCA IN MEDICINA MOLECOLARE

CICLO XXX

INFARTO MIOCARDICO GIOVANILE E VARIANTI ESOMICHE CHE INFLUENZANO I LIVELLI PLASMATICI DI COLESTEROLO

Coordinatore:

Chiar.ma Prof.ssa Stefania Conti

Tutore:

Chiar.mo Prof. Diego Ardissino

Dottorando: Dott.ssa Chiara Manzalini

Anni 2014/2018

INDICE

INTRODUZIONE.....	pag. 1
1. LA CARDIOPATIA ISCHEMICA	pag. 1
1.1. Epidemiologia.....	pag. 1
1.2. Fisiopatologia	pag. 2
1.3. Fattori di rischio e prevenzione	pag. 3
2. IL COLESTEROLO.....	pag. 4
2.1. Livelli di regolazione della biosintesi del colesterolo	pag. 4
2.2. Meccanismi di trasporto del colesterolo	pag. 5
3. L'IPERCOLESTEROLEMIA FAMILIARE	pag. 7
3.1. Alterazione della via di segnalazione dell'LDLR e ipercolesterolemia	pag. 7
3.2. Il gene <i>LDLR</i>	pag. 9
3.3. Il gene <i>APOB</i>	pag. 10
3.4. Il gene <i>PCSK9</i>	pag. 11
3.5. Il gene <i>LDLRAP1</i>	pag. 12
4. GENETICA DELLA CARDIOPATIA ISCHEMICA.....	pag. 13
4.1. Basi di genetica molecolare umana: disordini rari	pag. 14
4.2. L'era dei <i>Genome Wide Association Studies</i> : malattie comuni complesse	pag. 15
5. IL SEQUENZIAMENTO DI NUOVA GENERAZIONE: <i>NEXT</i> <i>GENERATION SEQUENCING</i>	pag. 16
5.1. <i>Next Generation Sequencing</i> e malattie mendeliane	pag. 17
5.2. <i>Next Generation Sequencing</i> e malattie comuni e complesse	pag. 18
6. NUOVI GENI CORRELATI AL METABOLISMO DEL COLESTEROLO	pag. 19
6.1. I geni <i>ABCG5/ABCG8</i> e il gene <i>NPC1L1</i>	pag. 19
6.2. Il gene <i>ABO</i>	pag. 20
6.3. Il gene <i>MYLIP</i>	pag. 20
6.4. Il gene <i>CYP7A1</i>	pag. 21
6.5. Il gene <i>INSIG2</i>	pag. 21
6.6. Il gene <i>LIPA</i>	pag. 22
6.7. Il gene <i>TM6SF2</i>	pag. 22
6.8. Il gene <i>APOE</i>	pag. 23

6.9. Il gene <i>ANGPTL3</i>	pag. 23
7. IL CONCETTO DI <i>MISSING HERITABILITY</i>	pag. 24
SCOPO DELLA TESI.....	pag. 26
MATERIALI E METODI.....	pag. 27
1. POPOLAZIONE DELLO STUDIO E RACCOLTA DEI CAMPIONI	pag. 27
2. ANALISI MOLECOLARE.....	pag. 27
3. STUDIO E SELEZIONE DEI GENI	pag. 28
4. STUDIO DELLE VARIANTI RARE.....	pag. 28
5. DETERMINAZIONE DEI LIVELLI PLASMATICI DI COLESTEROLO E VALUTAZIONE DELLE STENOSI CORONARICHE	pag. 29
6. ANALISI STATISTICA	pag. 29
RISULTATI.....	pag. 30
1. SEQUENZIAMENTO E SELEZIONE DEI GENI	pag. 30
2. EFFETTO DELLE VARIANTI GENETICHE SUL RISCHIO DI INFARTO MIOCARDICO	pag. 30
3. EFFETTO DELLE VARIANTI GENETICHE SUL FENOTIPO INTERMEDIO	pag. 31
DISCUSSIONE	pag. 32
TABELLE E FIGURE	pag. 35
BIBLIOGRAFIA.....	pag. 55

INTRODUZIONE

1. LA CARDIOPATIA ISCHEMICA

1.1 Epidemiologia

La cardiopatia ischemica, che comprende l'angina pectoris stabile, le sindromi coronariche acute e la morte cardiaca improvvisa, rappresenta attualmente la principale causa di mortalità e morbidità al mondo, e lo sarà fino al 2050 (1).

In Italia, dalla metà degli anni '70, si è assistito ad un *trend* in discesa della mortalità per cardiopatia ischemica. In particolare, tra il 1980 e il 2000, questo decremento è stato per lo più attribuito alle procedure terapeutiche in fase acuta, alla prevenzione primaria, secondaria e ad un cambiamento dei fattori di rischio nella popolazione dovuti alla modificazione dello stile di vita (2). Dal 1980 i tassi di mortalità negli uomini, aggiustati per età, sono diminuiti da 267,1 a 141,3 per 100.000 abitanti; analogamente, nelle donne di età compresa tra i 25 anni e gli 84 anni, i tassi di mortalità aggiustati per età sono diminuiti da 161,3 a 78,8. Questa decrescita ha comportato nel 2000 42.930 morti coronariche in meno (24.955 negli uomini e 17.975 nelle donne). Circa il 40% di questa diminuzione è associata all'esistenza, al giorno d'oggi, di specifici trattamenti per lo scompenso cardiaco (14%) e alle terapie di prevenzione secondaria dopo un infarto miocardico o una rivascolarizzazione (6%). Circa il 55% è invece associato ai cambiamenti nei più importanti fattori di rischio cardiovascolare nella popolazione italiana, in particolare la riduzione della pressione arteriosa (25%) e della colesterolemia totale. Ciononostante la cardiopatia ischemica rimane tutt'ora la principale causa di mortalità in Italia, rappresentando il 13% della mortalità generale e il 32% delle morti associate al sistema circolatorio (3). La più grave manifestazione della cardiopatia ischemica è rappresentata dall'infarto miocardico (STEMI), classicamente definito come conseguenza della morte di cellule miocardiche successivamente ad una prolungata ischemia; ogni anno in Italia il numero di pazienti ricoverati per infarto miocardico risulta essere di 180 mila (4).

1.2 Fisiopatologia

L'omeostasi del sistema cardiovascolare è governato dalla stretta interazione tra fattori genetici e ambientali; l'alterazione di quest'ultima è responsabile di processi patogenetici tipici delle malattie complesse quali l'aterosclerosi, e le sue complicazioni, l'infarto miocardico e l'ictus (5). Essendo l'aterosclerosi la principale causa della cardiopatia ischemica, questo fenomeno verrà descritto con maggiore attenzione.

Il più importante centro di regolazione dell'omeostasi del sistema cardiovascolare è rappresentato dall'endotelio dei vasi sanguigni, il quale è formato da un monostato di cellule che riveste l'intero sistema. Se in condizione fisiologiche l'endotelio vascolare è in grado di sostenere le funzioni dei vari tessuti e organi dell'organismo, questo non avviene in condizioni patologiche, dove la sua disfunzione provoca la manifestazione locale e sistemica delle malattie cardiovascolari aterosclerotiche (5).

La disfunzione delle cellule endoteliali, in aree prone allo sviluppo di lesioni della vascolatura arteriosa, si manifesta in cambiamenti precoci facilmente rilevabili durante il loro sviluppo, quali la formazione di placche focali e l'intrappolamento e modificazione chimico-fisica di particelle lipoproteiche nello spazio sub-endoteliale (5, 6, 7, 8). Nel corso di questi eventi si osservano ulteriori e complessi processi patologici, tra i quali la formazione iniziale di cellule schiumose (*foam cells*) e il successivo sviluppo della placca fibromuscolare. Le cellule schiumose, che contraddistinguono le lesioni aterosclerotiche precoci, si sviluppano in seguito alla selettiva transizione dei monociti circolanti dal sangue all'intima, dove differenziando in macrofagi internalizzano lipoproteine modificate; questa fase rappresenta uno stadio infiammatorio ancora reversibile nella formazione della placca aterosclerotica, poiché si tratta di un semplice accumulo di lipidi liberi o sottoforma di cellule schiumose. La placca fibromuscolare si sviluppa invece in seguito all'attivazione endoteliale, e quindi all'attivazione dalle molteplici chemochine, fattori di crescita e macrofagi, i quali agiscono sulle cellule prossime della muscolatura liscia inducendo sia la loro proliferazione sia la sintesi di componenti della matrice extracellulare nel compartimento intimale (5). Il progressivo rimodellamento strutturale di queste lesioni in via di sviluppo sono causa della formazione di un tappo di fibrina che ricopre un nucleo necrotico ricco di lipidi, lipoproteine ossidate, cristalli di colesterolo e detriti cellulari; si osserva inoltre rimodellamento e calcificazione della matrice. Le estremità di tali placche contengono anche una ricca popolazione di cellule infiammatorie (cellule T e macrofagi attivati, cellule *natural killer* T e cellule dendritiche) che hanno il potenziale di modulare il fenotipo pro-infiammatorio endoteliale; di conseguenza

anch'esse contribuiscono all'instabilità della placca mediante modificazioni proteolitiche delle componenti della matrice. Se si parla di placche vulnerabili o instabili, quello che si osserva in queste condizioni è una transizione catastrofica della lesione aterosclerotica: rottura della placca stessa e rilascio dell'intero contenuto trombogenico del nucleo necrotico con occlusione trombotica (5). In caso contrario si può assistere ad erosioni intime superficiali senza rottura della placca (9). Ciononostante l'attivazione di processi apoptotici delle cellule endoteliali, con denudamento dell'endotelio, rappresenta una fase critica per la formazioni di eventuali trombi (10). Generalmente queste erosioni superficiali si verificano sulla superficie delle lesioni contenenti abbondanti cellule muscolari lisce e proteoglicani e pochi macrofagi, e sono tipicamente associate a regioni con un flusso sanguigno disturbato. Possono ugualmente instaurarsi lesioni ancora più stabili. Queste sono caratterizzate da un folto tappo fibroso con scarso contenuto di lipidi e cellule infiammatorie, il quale può progressivamente invadere il lume del vaso causando sintomi ischemici, senza però necessariamente concludersi con eventi trombotici. Degno di nota è che in un dato individuo possono coesistere placche aterosclerotiche multiple, ciascuna in una propria fase evolutiva patobiologica (5).

1.3 Fattori di rischio e prevenzione

Negli ultimi decenni sono stati fatti grandi progressi nella prevenzione delle malattie cardiovascolari grazie, principalmente, ad una modificazione delle loro cause. Inoltre, si ritiene che la prevenzione secondaria di queste malattie, avvalendosi di una terapia medica aggressiva, ridurrà in modo drastico la loro ricorrenza in soggetti nei quali si sono già conclamati episodi coronarici acuti (11). *L'American Heart Association (AHA)* e *l'American College of Cardiology (ACC)* hanno pubblicato importanti raccomandazioni per l'intervento medico in pazienti con malattie cardiovascolari e altre forme di malattie aterosclerotiche (12). Analogamente esiste la possibilità di ridurre il rischio di cardiopatia ischemica in pazienti che non hanno mai riscontrato episodi coronarici mediante una prevenzione di tipo primario. Tuttavia, questi soggetti hanno un rischio di malattia estremamente variabile, ragione per la quale è stato necessario determinare dei *range* di intensità nei quali intervenire. Infatti, un'efficace prevenzione primaria richiede, in primo luogo, una valutazione del rischio al fine di classificare i pazienti e selezionare coloro per quali sarà possibile intervenire in modo incisivo (11).

I più importanti fattori di rischio delle malattie cardiovascolari ad oggi noti sono sei:

1. Familiarità per cardiopatia ischemica;
2. Ipercolesterolemia;
3. Ipertensione arteriosa;
4. Indice di massa corporea elevato
5. Diabete Mellito
6. Fumo di sigaretta

La relazione quantitativa tra questi fattori di rischio e il rischio di cardiopatia ischemica è stata illustrata dal *Framingham Heart Study* e molti altri lavori, i quali hanno dimostrato che i più importanti fattori di rischio sono additivi nel potere predittivo di una malattia (11, 13). Di conseguenza, il rischio totale di una persona può essere stimato dalla totalità del rischio impartito da ciascuno di questi maggiori fattori.

2. IL COLESTEROLO

2.1 Livelli di regolazione della biosintesi del colesterolo

Il colesterolo, ad oggi riconosciuto tra i più importanti fattori di rischio della cardiopatia ischemica, può essere sintetizzato dall'organismo o assunto con la dieta. Un adulto a dieta povera di colesterolo sintetizza di regola circa 800 mg di colesterolo al giorno. Nei mammiferi il principale sito di sintesi del colesterolo è il fegato, sebbene anche l'intestino ne formi significative quantità (14).

La velocità con la quale il colesterolo può essere sintetizzato nel fegato e nell'intestino è strettamente dipendente dalla concentrazione intracellulare del colesterolo stesso in questi organi. In particolare si parla di un meccanismo di regolazione a *feedback*, il quale è mediato dalla variazione della quantità e dell'attività della 3-idrossi-3metilglutaril CoA reduttasi (HMG-CoA); quest'ultimo infatti è un enzima che catalizza la formazione del mevalonato, tappa di comando della biosintesi del colesterolo (14).

I meccanismi di regolazione dell'HMG-CoA reduttasi possono essere diversi. In primo luogo è importante osservare che la velocità di sintesi dell'mRNA per la HMG-CoA reduttasi è regolata da una importante proteina, nota come SREBP (*sterol regulatory element-binding protein*), la quale lega l'elemento di regolazione degli steroli. Questa proteina, nel suo stato

inattivo, è localizzata nella membrana del reticolo endoplasmatico, dove è associata a SCAP (*SREBP-cleavage-activating protein*), proteina integrale di membrana e sensore del colesterolo (14). Quando i livelli del colesterolo diminuiscono, la proteina SCAP accompagna SREBP in piccole vescicole membranose all'apparato di Golgi, dove subisce due importanti tagli proteolitici sequenziali ad opera di una serina proteasi e di una metalloproteasi. Grazie a quest'ultimi, il dominio di legame del DNA della proteina SREBP può quindi migrare a livello nucleare e regolare la trascrizione di diversi geni, tra i quali il gene dell' HMG-CoA reduttasi (Figura 1). Anche la traduzione del trascritto per l' HMG-CoA reduttasi è finemente regolato; infatti questa viene inibita dai metaboliti non sterolici, derivati dal mevalonato o dal colesterolo assunto con la dieta. Infine è importante osservare che in presenza di un aumento della concentrazione di composti sterolici, come il colesterolo stesso, si osserva una maggiore suscettibilità dell'HMG-CoA reduttasi a degradazione proteolitica; mentre la sua fosforilazione si traduce in una ridotta attività enzimatica dell' HMG-CoA reduttasi (14).

2.2 Meccanismi di trasporto del colesterolo

Il colesterolo, così come i trigliceridi, vengono trasportati nei liquidi corporei sotto forma di particelle lipoproteiche, ovvero molecole composte da un nucleo di lipidi idrofobici circondato da un involucro di lipidi più polari e proteine (14). Quest'ultime sono note come apoproteine, le cui funzioni sono rendere solubili i lipidi idrofobici e contenere il segnale di riconoscimento cellulare. Le particelle lipoproteiche, in funzione della loro densità, vengono classificate in chilomicroni, lipoproteine a densità molto bassa (VLDL), lipoproteine a densità intermedia (IDL), lipoproteine a bassa densità (LDL) e lipoproteine ad alta densità (HDL) (Tabella 1). Elevate concentrazioni di colesterolo nel sangue rappresentano la principale causa nella formazione di placche aterosclerotiche nelle arterie dell'organismo, e tale eccesso è presente sotto forma di particelle LDL; questa è la ragione per la quale il colesterolo LDL generalmente è identificato con il termine "colesterolo cattivo" (15). Al contrario le HDL vengono chiamate "colesterolo buono" poiché aiutano a rimuovere il colesterolo dai vasi sanguigni. Infatti le HDL legano ed esterificano il colesterolo proveniente dai tessuti periferici e lo trasportano al fegato, che lo elimina nel lume intestinale sotto forma di sali biliari o colesterolo libero, o ad organi steroideogenici che lo utilizzano per formare gli ormoni steroidei. I chilomicroni, le particelle lipoproteiche più grandi e costituite per il 90% da trigliceridi, rappresentano la principale forma di trasferimento del colesterolo, e degli altri grassi ottenuti con la dieta, dall'intestino ai vari distretti dell'organismo. I trigliceridi dei chilomicroni vengono rilasciati a

livello tissutale dove subiscono idrolisi catalizzata da lipasi lipoproteiche, enzimi localizzati sul rivestimento dei vasi sanguigni di determinati tessuti, che utilizzano gli acidi grassi sia come sostanze nutrienti sia per la sintesi dei lipidi; infine il fegato ingloba le rimanenze dei chilomicroni, ricche di colesterolo (14).

Al contrario, in presenza di un eccesso epatico di colesterolo, e altri lipidi, questi vengono esportati dal fegato sotto forma di VLDL dove trasferiscono il loro contenuto in trigliceridi al tessuto adiposo e ad altri tessuti periferici. Le VLDL vengono quindi convertite prima in IDL e successivamente a LDL. Gli epatociti e le cellule dei tessuti periferici captano quindi le LDL mediante un meccanismo di endocitosi-recettore mediata, le cui tappe sono le seguenti:

- L'apolipoproteina B-100, localizzata sulla superficie di una particella di LDL, si lega al suo specifico recettore LDLR (*low-density lipoprotein receptor*), il quale è localizzato in speciali zone della membrana ricche di clatrina;
- A seguito del legame dell'LDL con il recettore LDLR, avviene internalizzazione cellulare complesso LDLR/LDL per endocitosi recettore-mediata;
- Le vescicole contenenti le LDL si fondono successivamente con i lisosomi contenenti enzimi degradativi; avviene quindi idrolisi sia della componente proteica delle LDL sia degli ester del colesterolo. Al contrario il recettore delle LDL ritorna immutato sulla membrana plasmatica;
- Il colesterolo non esterificato rilasciato può essere utilizzato per la biosintesi delle membrane o essere riesterificato per essere immagazzinato all'interno della cellula. La riesterificazione del colesterolo è essenziale, in quanto elevate concentrazioni di colesterolo non esterificato possono danneggiare le membrane cellulari.

È importante osservare che quando il colesterolo è abbondante a livello cellulare, non vengono più sintetizzati recettori per le LDL, e quindi viene bloccata l'internalizzazione di altro colesterolo mediata dall'endocitosi delle LDL plasmatiche (meccanismo di regolazione a *feedback*). Ciononostante sono diverse le patologie ad oggi riconosciute responsabili di elevati livelli plasmatici di LDL e, se vengono considerate solo quelle associate ad un mancato e scorretto funzionamento di LDLR, si parla di sitosterolomia, ipercolesterolemia autosomica recessiva, disordine familiare difettivo per l'apolipoproteina e ipercolesterolemia familiare (FH). Quest'ultima è stato il primo disordine monogenico riconosciuto essere responsabile di elevati livelli plasmatici di colesterolo LDL (14).

3. L'IPERCOLESTEROLEMIA FAMILIARE

L'ipercolesterolemia familiare (*Familial Hypercholesterolemia/FH*) è una comune patologia monogenica caratterizzata da elevate concentrazioni plasmatiche di colesterolo LDL ed aterosclerosi precoce, ragione per la quale questo disordine è causa di una significativa morbidità e mortalità. Dal punto di vista clinico la FH si manifesta con xantoma tendinei, xantelasma, arco corneale prematuro o semplicemente con elevati livelli di colesterolo LDL senza altre evidenze cliniche (16).

La forma omozigote dell'ipercolesterolemia familiare (HoFH) è estremamente rara, manifestandosi con una frequenza di 1:1.000.000 (o 1:100.000 in alcune popolazioni libanesi). Al contrario l'ipercolesterolemia familiare eterozigote (HeFH) è molto più diffusa interessando 1 su 67-500 individui, indipendentemente dalla zona geografia considerata o dall'etnia di appartenenza; un esempio è rappresentato dagli ebrei sudafricani ashkenazi e i libanesi cristiani la cui prevalenza è pressoché paragonabile essendo rispettivamente 1:67 e 1:85 (16).

Considerando che <1% delle popolazioni dei diversi paesi si sottopone a specifici esami di *screening* per la FH, è stato valutato come questa malattia sia generalmente sotto-diagnosticata; unica eccezione è l'Olanda, dove il 71% della popolazione si sottopone a tali esami (34).

Degno di nota è che esistono condizioni fenotipicamente simili alla FH, quali la xantomatosi cerebro tendinea (CTX), la malattia da accumulo di esteri del colesterolo e la sitosterolemia, ragione per la quale una diagnosi differenziale si rivela fondamentale (16, 17). Infatti un intervento terapeutico efficace e una appropriata consulenza genetica sono successive solo ad una corretta diagnosi della patologia, la quale è possibile dopo accurata osservazione clinica, quantificazione del colesterolo LDL e test genetici di conferma (16).

3.1 Alterazione della via di segnalazione dell'LDLR e ipercolesterolemia

L'LDLR ha la funzione di rimuovere le particelle di LDL dal plasma mediante l'apolipoproteina B (APOB). In particolare a seguito del loro legame, avviene internalizzazione del complesso LDL/LDLR/APOB negli epatociti mediante invaginazione della membrana plasmatica in regioni ricche di clatrina, e formazione di endosomi grazie alla proteina LDLRAP1 (*low density lipoprotein receptor adaptor protein 1*) (Figura 2). All'interno dell'endosoma avviene la separazione del complesso LDL/LDLR; quest'ultimo può quindi andare incontro a due diversi destini: degradazione lisosomiale mediata dalla proteina PCSK9

(*Proprotein convertase subtilisin/kexin type 9*) o trasporto sulla membrana plasmatica cellulare per essere riciclato. In alternativa LDLR può legare PCSK9 quando è ancora localizzato sulla membrana plasmatica, legame il cui risultato finale è sempre l'internalizzazione lisosomiale di LDLR e la sua degradazione (18).

È evidente, quindi, come in seguito ad alterazioni funzionali o quantitative a carico di proteine coinvolte nella via di segnalazione dell'LDL possa verificarsi una scorretta rimozione dell'LDL plasmatico e l'instaurarsi un quadro ipercolesterolemico. Un esempio è rappresentato da soggetti con FH dovuta a mutazioni nulle (ovvero con perdita di funzione della proteina) nel gene *LDLR*; in questi soggetti si osserva una riduzione della proteina LDLR pari a circa il 50% rispetto a soggetti sani. Analogamente soggetti omozigoti (o eterozigoti composti), ovvero con entrambi gli alleli mutati per il gene *LDLR*, presentano un severo quadro patologico; infatti questi soggetti sono totalmente incapaci di produrre il recettore LDLR o di produrre anche piccole dosi di recettore correttamente funzionante. La via di segnalazione dell'LDLR può essere alterata anche in seguito a mutazioni a carico di altri geni. Ad esempio, mutazioni nel gene *APOB* possono alterare il legame della proteina da esso codificata con il recettore LDLR, compromettendo il legame dello stesso colesterolo LDL con il recettore. Un secondo esempio è rappresentato da mutazioni che *up*-regolano il gene PCSK9, causando una iperfunzione della proteina da esso codificata, aumentata degradazione dell'LDLR e l'instaurarsi di un quadro ipercolesterolemico. È importante osservare che ciò che accomuna queste mutazioni, associate ad un *pattern* di ereditarietà semidominante, è un effetto di dosaggio genico paragonabile e l'instaurarsi di un fenotipo clinico finale molto simile. Infine in presenza di entrambi gli alleli del gene *LDLRAP1* mutati si instaura un fenotipo clinico che riproduce la FH omozigote, in accordo con un'ereditarietà autosomica recessiva (16).

Recenti studi hanno suggerito che la presenza di più mutazioni in geni codificanti proteine coinvolte in questa via di segnalazione sono responsabili di un effetto finale additivo, il quale può essere intermedio a quello osservato in individui eterozigoti ed omozigoti. Un fenotipo con mutazioni multiple può implicare qualsiasi combinazione di mutazioni nei geni *APOB*, *LDLR*, *PCSK9* e *LDLRAP1*; tuttavia sebbene il fenotipo derivato da queste sia apparentemente simile, in realtà può differire sia per la severità del quadro clinico con la quale si manifesta la patologia sia nei livelli di colesterolo LDL (19, 20). Negli omozigoti, i più elevati livelli di colesterolo LDL sono stati osservati in pazienti incapaci di produrre il recettore per l'LDL, seguono a questi soggetti con un'alterata funzione dell'LDLR o soggetti con ipercolesterolemia autosomica recessiva (ARH), e infine individui con mutazioni *gain-of function* per *APOB* o *PCSK9* (19). Un interessante studio francese ha riportato che mutazioni nei geni *LDLR* e

PCSK9 aumentano in modo paragonabile i livelli di colesterolo LDL; al contrario mutazioni nel gene *APOB* sono associate ad un aumento meno severo dei livelli di colesterolo LDL (21, 22). Dal punto di vista eziologico, quindi, esistono forme di FH sia poligeniche sia monogeniche, sebbene in quest'ultimo caso siano stati identificati un limitato numero di geni. Inoltre, l'identificazione di diversi casi di FH non geneticamente spiegati ha supposto l'esistenza di *loci* non ancora caratterizzati (23).

3.2 Il gene *LDLR*

Il gene *LDLR* è un gene di 45Kb costituito da 18 esoni e 17 introni; la proteina codificata da questo gene si compone di 860 aminoacidi, comprensivi di un peptide segnale di 21 residui aminoacidici (16).

Mutazioni nel gene *LDLR* rappresentano la più importante causa di FH monogenica, riducendo quantitativamente l'espressione di questa proteina o compromettendo la sua capacità funzionale di rimozione del colesterolo LDL plasmatico (24, 25). Ad oggi sono circa 1600 le mutazioni note a carico di questo gene, come riportato dai *database Human Gene Mutation* (26), *ClinVar* (27) e lo *UK British Heart Foundation* (28). Tuttavia è stata effettuata una sottostima relativamente alle mutazioni nel gene *LDLR* individuate in pazienti con FH, ragione per la quale nella pratica clinica non sempre queste varianti sono state riportate. La gamma di mutazioni a carico di questo gene è estremamente variabile e include delezioni, duplicazioni, mutazioni *missenso*, di *splicing*, nulle, mutazioni nella regione promotrice e molte altre. Degne di nota sono le varianti genetiche che interessano l'esone 4, poiché questa regione codifica per una porzione essenziale per il legame dell'LDL con APOB, e sembrano quindi essere causa del fenotipo più grave derivante dalla compromissione di questo gene (16). In funzione del loro effetto sul fenotipo finale sono state identificate cinque classi di mutazioni a carico del gene *LDLR*:

1. Mutazioni nulle, ovvero responsabili di una perdita di funzione della proteina. Queste generalmente interessano la regione promotrice, introducono codoni di stop prematuri o alterano la *reading frame* del gene con conseguente non traduzione della proteina o traduzione di una proteina non funzionante (25, 29)
2. Mutazioni responsabili di difetti post-traduzionali, le quali alterano la corretta traslocazione di LDLR sulla membrana cellulare (25, 29, 30)
3. Mutazioni responsabili di un'alterata capacità di LDLR di legare efficacemente il proprio substrato (25, 29)

4. Mutazioni che sebbene non alterino la capacità di LDLR di legare il proprio substrato, riducono la sua capacità endocitosi recettore mediata (25, 29)
5. Mutazioni responsabili di un difettoso riciclo dell'LDLR (25, 29)

È importante osservare che non tutte le mutazioni a carico del gene *LDLR* sono di facile interpretazione, come avviene invece per le mutazioni *missenso* o varianti genetiche che interessano la regione promotrice. Questa è la ragione per la quale studi funzionali, quali le analisi di segregazione, si rivelano di grande interesse. Ad ogni modo, una corretta e accurata recensione di tutte le varianti genetiche benigne o incerte individuate è fondamentale sia dal punto di vista genetico sia clinico.

3.3 Il gene *APOB*

Il gene *APOB* ha una dimensione di 42 kb, ed è composto da 29 esoni e 28 introni che possono dare origine a due diverse isoforme della proteina. L'isoforma *APOB*-100, il più importante componente proteico delle particelle di colesterolo LDL, ha una lunghezza di 4563 amminoacidi organizzati in quattro diversi domini, ed è espresso principalmente dagli epatociti; l'isoforma *APOB*-48 è invece espressa nell'intestino tenue (16). Il gene *APOB* è organizzato con la maggior parte dei 29 esoni localizzati all'estremità 5' del gene e nella prima porzione di questo (31, 32).

Difetti nella proteina *APOB*-100 sono causa di un fenotipo clinico simile alla FH, noto con il termine ipercolesterolemia familiare difettiva per *APOB*-100 (FDB). Individui FDB, in assenza di specifici test genetici, possono essere clinicamente indistinguibili da pazienti con altre forme di ipercolesterolemia familiare (31). Mutazioni in *APOB* seguono una ereditarietà autosomica semidominante con penetranza incompleta, e la maggior parte di queste mutazioni si presentano in eterozigosi (33). Al contrario del gene *LDLR*, sono limitate le mutazioni genetiche identificate a carico del gene *APOB* e la maggior parte di esse interessano principalmente l'esone 26. Inoltre, la più diffusa di queste mutazioni coinvolge l'amminoacido 3500 che è responsabile di un alterato *folding* proteico (34). Circa il 5% dei casi di ipercolesterolemia diagnosticati sono attribuiti alla FDB, con una prevalenza variabile da 1:500 a 1:1000 in funzione della popolazione studiata (34, 36). Ulteriori varianti patogenetiche a carico del gene *APOB* includono le mutazioni p.Arg3500Trp (spesso riscontrata in individui affetti da FH della Cina o Malesia), p.Arg3531Cys (con una riduzione del legame dell'LDL maggiore rispetto alla mutazione p.Arg3500Gln e una meno pronunciata ipercolesterolemia), e

la mutazione p.His3543Tyr (16). Quest'ultima è stata principalmente identificata nella popolazione tedesca con una frequenza quattro volte più alta rispetto alla mutazione p.Arg3500Gln (37).

Più recentemente sono state identificate anche le varianti genetiche p.Arg3059Cys e p.Lys3394Asn, riscontrate per lo più in famiglie con FH a penetranza variabile. Studi funzionali condotti *in vitro* hanno evidenziato come quest'ultime siano responsabili di una debole capacità di legame dell'LDLR (38). Ulteriori studi hanno permesso l'identificazione anche della mutazione p.Arg50Trp, la quale vede come protagonista principale l'esone 3 (39).

Recenti test clinici a carico del gene *APOB* sono principalmente focalizzati sull'esone 26 e la mutazione p.Arg3500Trp. A seguito delle diverse alterazioni genetiche che sono state identificate è sicuramente importante osservare come il sequenziamento dell'intero gene *APOB* sia di notevole importanza dal punto di vista clinico al fine di comprendere le causa di determinati fenotipi.

3.4 Il gene PCSK9

Il gene *PCSK9* ha una dimensione di 25 kb ed è composto da 12 esoni. Questo gene traduce per una serina proteasi di 692 aminoacidi, i quali sono organizzati in una sequenza segnale di 30 aminoacidi seguita da tre diversi domini: un prodominio, un dominio catalitico, e un dominio C-terminale (16).

PCSK9, nono membro nei mammiferi della famiglia di proproteine convertasi, è coinvolta nell'attivazione, inattivazione e regolazione di proteine secrete e nella localizzazione cellulare. In particolare, *PCSK9* gioca un ruolo cruciale nel metabolismo del colesterolo, poiché controlla i livelli cellulari del recettore LDLR e di conseguenza anche i livelli del colesterolo plasmatico LDL. Queste osservazioni hanno fatto in modo che il gene *PCSK9* fosse protagonista di innumerevoli studi, ricoprendo un elevato successo (16).

Mutazioni a carico del gene *PCSK9*, responsabili dell'ipercolesterolemia autosomica dominante, sono estremamente rare, contribuendo con una percentuale decisamente inferiore rispetto a mutazioni a carico dei geni *LDLR* ed *APOB* (42). Grazie a questi iniziali studi condotti sul gene *PCSK9*, che hanno evidenziato come mutazioni *gain-of-function* fossero responsabili dell'ipercolesterolemia autosomica dominante, sono state successivamente descritte numerose mutazioni *loss-of-function* associate ad una riduzione dei livelli di colesterolo plasmatico. Un importante studio prospettico ha individuato due interessanti mutazioni non senso in *PCSK9*, note come Y142X e C679X, presenti nel 2,6% della

popolazione afro-americana; queste mutazioni risultano associate ad una riduzione del 28% del colesterolo plasmatico medio LDL e ad una riduzione dell'88% del rischio di cardiopatia ischemica. Lo stesso studio ha dimostrato un'importante associazione tra la variante R46L nei Caucasici (individuata con una frequenza del 3,2%) ed una riduzione del rischio di cardiopatia ischemica del 47%, sebbene si sia osservata una riduzione dei livelli plasmatici di colesterolo LDL del solo 15% (41).

L'individuazione dell'associazione tra mutazioni *loss-of-function* a carico del gene *PCSK9* e la riduzione del rischio di cardiopatia ischemica, mediata da una variazione dei livelli del colesterolo plasmatico LDL, ha fatto in modo che *PCSK9* rappresentasse al giorno d'oggi il principale *target* di farmaci di ultima generazione (anticorpi monoclonali) utilizzati per abbassare il colesterolo LDL (16, 40, 42). I primi anticorpi monoclonali anti-*PCSK9* sviluppati e tutt'ora sottoposti a trial clinici, noti come alirocumab e evolocumab, hanno dimostrato di ridurre le concentrazioni di colesterolo plasmatico LDL del 55-60% rispetto al placebo, sia quando somministrati singolarmente sia quando somministrati in combinazione a statine. Gli anticorpi anti-*PCSK9* sono estremamente vantaggiosi non solo perché ben tollerati dall'organismo, ma anche perché l'unica reazione avversa derivante dal loro utilizzo è una marginale reazione nel sito di iniezione. Ciononostante, attualmente il loro utilizzo è ampiamente limitato dagli elevati costi (43).

3.5 Il gene *LDLRAP1*

Varianti genetiche associate al gene *LDLRAP1* sono causa dell'ipercolesterolemia autosomica recessiva (*ARH/ autosomal recessive hypercholesterolemia*) che segue un *pattern* di ereditarietà autosomico recessivo, ovvero la patologia si manifesta nei soli pazienti omozigoti o eterozigoti composti (16). Per questa ragione gli eterozigoti, sebbene possano avere livelli di colesterolo LDL più elevati rispetto ai non portatori, non presentano ipercolesterolemia, (44).

Il gene *LDLRAP1* contiene 9 esoni e 8 introni, e codifica per un prodotto proteico di 308 aminoacidi (44, 45). La proteina *LDLRAP1* è coinvolta nella formazione di vescicole ricche di clatrina e nell'endocitosi del complesso LDL/LDLR (46, 47). Mutazioni in *LDLRAP1* responsabili della *ARH* sono estremamente rare, infatti è stimata una loro frequenza in <1 su 5.000.000 individui (35). In Italia la *ARH* ha una incidenza maggiore in Sardegna (1:40.000) a causa della isolata posizione geografica di quest'isola e di una maggiore consanguineità dei suoi abitanti (48).

Dal punto di vista clinico pazienti affetti da ARH hanno caratteristiche simili a soggetti con HoFH, con solo alcune differenze: pazienti con ARH hanno livelli di colesterolo HDL notevolmente maggiori rispetto a soggetti con HoFH, rara stenosi critica e ridotta progressione della malattia nella radice aortica (44, 45). Diversamente dai pazienti con HoFH, in cui la malattia si presenta precocemente, l'età nella quale si presenta la ARH varia molto tra la prima e la quinta decade di vita; inoltre pazienti che manifestano ARH tardivamente tendono ad avere un quadro clinico che riproduce molto la HeFH (47, 49). Inoltre soggetti con ARH rispondono efficacemente alle terapie volte ad una riduzione del quadro lipidico (50).

4. GENETICA DELLA CARDIOPATIA ISCHEMICA

La cardiopatia ischemica, ed in particolare l'infarto miocardico, è una malattia multifattoriale la cui eziopatogenesi deriva dall'interazione tra fattori di rischio ambientali e fattori di rischio genetici. Tuttavia è importante osservare come questa interazione vari a seconda dell'età di insorgenza della patologia stessa. Infatti, tanto più precocemente si verifica un infarto miocardico, tanto maggiore sarà il contributo dei fattori genetici alla patologia; al contrario se l'infarto occorre tardivamente il contributo dei fattori genetici a tale evento tende ad affievolirsi, mentre accresce il contributo dei fattori ambientali alla patologia (Figura 3). In questo senso l'infarto miocardico giovanile rappresenta sicuramente un modello sperimentale di elevato interesse al fine di comprendere l'entità del contributo genetico in questa malattia.

La familiarità per cardiopatia ischemica, inoltre, è ormai riconosciuto essere un fattore di rischio indipendente per lo sviluppo di infarto miocardico, come dimostrato da studi condotti sui gemelli mono e dizigoti nati tra il 1886 e il 1990 ed utilizzando dati ricavati dallo *Swedish Twin Registry* (51, 52, 53). In particolare è stato osservato che la morte cardiaca in età giovanile di uno dei due gemelli monozigoti comporta un rischio notevolmente maggiore di andare incontro al medesimo evento per il gemello sopravvissuto rispetto al caso in cui nessuno dei due gemelli muoia per morte cardiaca. Al contrario, nel caso di gemelli dizigoti, e che quindi condividono metà del patrimonio genetico, il rischio è significativamente inferiore (54).

4.1 Basi di genetica molecolare umana: i disordini rari

Fino a poco tempo fa la genetica aveva concentrato i propri sforzi solo sulle malattie a carattere mendeliano: disordini rari a singolo gene che si verificano in meno dell'1% della popolazione (55). Queste malattie sono caratterizzate da mutazioni genetiche altamente penetranti che interessano, nella maggior parte dei casi, la porzione codificante del gene e che distruggono la funzione della proteina codificata determinando la comparsa di un fenotipo facilmente riconoscibile. Tali mutazioni si definiscono necessarie e sufficienti per l'instaurarsi del fenotipo "malattia". Tuttavia, inizialmente erano stati identificati un limitato numero di geni associati a queste malattie, poiché era richiesta la conoscenza preliminare della proteina coinvolta (55).

Il primo grande cambiamento è avvenuto nel 1980 con lo sviluppo di una tecnica genetica nota come clonaggio posizionale, la quale ha permesso di determinare, sebbene in modo approssimativo, la posizione cromosomica del gene responsabile della malattia senza conoscere in modo preliminare la proteina coinvolta. Secondo questa tecnica è possibile genotipizzare *pedigree*, da 2 a 3 generazioni, di soggetti affetti dalla malattia avvalendosi di marcatori di DNA con localizzazione cromosomica nota. In particolare, avvalendosi di studi di *linkage* è possibile confrontare la segregazione di questi marcatori tra individui malati e sani; solo la co-ereditarietà del marcatore negli individui malati è indicativo di come quest'ultimo sia fisicamente prossimo al gene coinvolto nella patologia (55). Inizialmente questi marcatori erano rappresentati da polimorfismi associati a frammenti di restrizione, le cui analisi mediante *Southern Blot* richiedevano 10-12 giorni; successivamente sono stati identificati marcatori maggiormente informativi, ovvero variabili ripetute di-, tri o -tetra nucleotidiche analizzabili in poche ore mediante reazioni di PCR (*Polymerase Chain Reaction*) (56). Grazie alla tecnica del clonaggio posizionale è stato quindi possibile individuare con minore difficoltà i geni responsabili di disordini rari a singolo gene. Dal punto di vista cardiovascolare, il primo di questi ad essere identificato è stato il gene per la catena β della miosina, il quale quando soggetto ad una specifica mutazione *missenso* è causa della cardiomiopatia ipertrofica familiare (57, 58, 59). Successivamente sono stati identificati molti altri geni responsabili di disordini cardiovascolari, quali la sindrome QT, la sindrome di Brugada, la fibrillazione atriale, la cardiomiopatia dilatativa, la sindrome *Wolff-Parkinson-White* e la displasia aritmogena del ventricolo destro (60-65). Inoltre, nel ventennio a seguire, per oltre 3.000 dei 7.000 disordini a singolo gene stimati, è stata identificata più di una variante genetica per singolo gene, così come è stato osservato nell'ipercolesterolemia familiare associata a mutazioni nel gene *LDLR*.

4.2 L'era dei *Genome Wide Association Studies*: malattie comuni complesse

Nonostante il grande successo degli ultimi vent'anni nello studio dei disordini rari a singolo gene, per quanto riguarda lo studio delle malattie comuni è stato necessario attendere maggiori sviluppi. Nel 1990 si è acquisita la consapevolezza dell'inadeguatezza degli studi di *linkage* per lo studio dei disordini comuni come le malattie coronariche o il cancro. Inoltre, diversi studi epidemiologici hanno dimostrato che la predisposizione genetica per le malattie comuni può essere il risultato della combinazione di multiple varianti genetiche, le quali singolarmente avrebbero solo un piccolo o moderato effetto sul fenotipo. Quindi, al contrario dei disordini rari, nelle malattie comuni la presenza di un singolo gene alterato non necessariamente è sufficiente o necessario per indurre il fenotipo malattia, il quale inoltre non segue un *pattern* di trasmissione mendeliano. Al fine di identificare i geni responsabili di questi disordini è stato necessario avvalersi di un differente approccio rispetto all'analisi di *linkage* (55). Infatti, l'applicazione di questo tipo di analisi allo studio dei disordini comuni richiederebbe centinaia di migliaia di marcatori separati tra loro da più di 6.000 basi, al contrario dei disordini rari a singolo gene dove questa analisi ne richiede solo alcune centinaia separati fino a 10 milioni di basi (66, 67). Lo studio dei disordini comuni quindi non si avvale né di studi di *linkage* né dello studio di famiglie, bensì di studi di associazione di migliaia di individui con (casi) o senza (controlli) la malattia e non correlati tra loro (68, 69). Tale approccio confronta la frequenza di specifici marcatori nei casi rispetto ai controlli che, se maggiormente rappresentati nei primi, sono indicativi del fatto che il marcatore è localizzato in una regione del DNA associata ad una maggiore predisposizione per la malattia (56).

Gli studi relativi alle malattie comuni sono iniziati nel 1990 e, data la mancata disponibilità di *array* genetici, questi sono stati inizialmente condotti partendo da una attenta selezione dei geni da analizzare in base alla loro funzione e confrontando la frequenza di mutazioni tra casi e controlli. Questo approccio ha permesso di individuare oltre 100 geni come possibili candidati ad una predisposizione genetica per le malattie cardiovascolari, ma quasi nessuno di essi è stato confermato nei successivi studi (70, 71).

La prima grande svolta è avvenuta nel 2005 con lo sviluppo di *microarray* inizialmente caratterizzati da 50.000 marcatori e successivamente ampliati fino a 1 milione; questi consistono di polimorfismi a singolo nucleotide (SNP) uniformemente distribuiti lungo il genoma (72). Grazie al progetto genomico *Hap Map* è stato possibile stabilire la localizzazione cromosomica di oltre 1 milione di questi SNP, e comprendere come mediamente il genoma umano ne contenga circa 3 milioni (73). Inoltre molteplici studi hanno indicato che, sebbene le

sequenze del genoma umano siano identiche al 99,5%, esistono circa 15 milioni di paia di basi che differiscono tra essi, ovvero 3 milioni di SNP che possono essere ritenuti responsabili della variabilità tra gli individui, compresa la suscettibilità ad una data malattia (74). L'elaborazione dei *microarray* ha permesso una massiva analisi dei dati e lo sviluppo di studi di associazione *genome-wide* (GWAS): approccio *unbiased*, che non richiede la conoscenza preliminare del marcatore o del gene candidato al rischio, e che permette di determinare sia la loro frequenza sia se sono maggiormente rappresentati nei casi rispetto ai controlli (72).

5. IL SEQUENZIAMENTO DI NUOVA GENERAZIONE: *NEXT GENERATION SEQUENCING*

Sebbene negli ultimi decenni siano stati effettuati considerevoli progressi per comprendere la correlazione genotipo-fenotipo, nel contesto delle malattie mendeliane e delle malattie comuni e complesse, ulteriori indagini devono ancora essere effettuate per comprendere appieno quest'associazione. Relativamente alle malattie a trasmissione mendeliana, l'obiettivo ultimo è quello di mappare tutte le malattie non ancora risolte; per quanto riguarda le malattie comuni e complesse, l'obiettivo è invece estendere lo studio sia a tutte le varianti genetiche a bassa frequenza (compresa tra 1:1000 e 1:20) sia alle varianti genetiche rare (frequenza < 1:1000). Con l'introduzione delle tecniche di nuova generazione, o NGS (*Next generation sequencing*), insieme ai progressi effettuati nelle tecniche di genotipizzazione, è stato possibile ottenere grandi risultati per il raggiungimento di questi obiettivi. Le tecniche NGS presentano diversi vantaggi: in primo luogo hanno permesso un abbattimento dei costi di sequenziamento del DNA rispetto al metodo Sanger; queste tecniche infatti effettuano il sequenziamento di miliardi di paia di basi in parallelo ed in tempi estremamente più ragionevoli. Inoltre, l'approccio di ibridazione sul quale si basano le piattaforme NGS hanno permesso l'analisi del solo "esoma", ovvero la porzione codificante del genoma. L'esoma costituisce l'1% di tutto il materiale genetico di un individuo ed è costituito da oltre 30 megabasi, circa 180.000 esoni. Il sequenziamento dell'esoma, noto come "*exome sequencing*", è giustificato dalla ricerca delle cause genetiche delle malattie rare ereditarie. Degno di nota, infatti, è che la maggior parte delle malattie a trasmissione mendeliana sono il risultato di mutazioni che interrompono la sequenza codificante di un gene; inoltre esiste un'ampia percentuale di mutazioni rare *missenso* predette essere deleterie per la funzione del gene stesso (75).

5.1 Next generation sequencing e malattie mendeliane

Le tecniche di sequenziamento di nuova generazione sono state determinanti per l'identificazione di nuovi geni associati alle malattie mendeliane di natura cardiovascolare. La maggior parte di questi studi sono stati effettuati grazie al sequenziamento dell'esoma di uno o pochi individui affetti da una specifica malattia; in questo caso le varianti genetiche individuate vengono confrontate con quelle di individui di riferimento (individui sani della stessa famiglia o non correlati tra loro). Solo le varianti presenti negli individui affetti, ma assenti nella popolazione sana di riferimento, sono state considerate possibili varianti causative la malattia (75). Questa strategia ha permesso l'associazione di varianti del gene *ANGPTL3* (*angiopoietin like 3*) all'ipolipidemia combinata familiare (76).

Tuttavia l'applicazione delle tecniche di *exome-sequencing* per la comprensione genetica di patologie a trasmissione mendeliana presenta ancora diversi limiti per le seguenti motivazioni:

- Non tutti i risultati negativi sono stati regolarmente riportati;
- la variante genetica causativa potrebbe non essere localizzata nell'esoma;
- la variante genetica causativa è localizzata nell'esoma, ma il gene di interesse potrebbe essere "sfuggito" al sequenziamento;
- il 5-10% dell' esoma è di difficile sequenziamento a causa dell'alto contenuto in GC (guanine e citosine);
- bassa penetranza della variante di interesse; questo significa che la variante causativa è presente in individui fenotipicamente affetti e non, e di conseguenza rischia di essere esclusa dall'analisi;
- la segregazione di un fenotipo in una famiglia può essere dovuta a fattori non genetici piuttosto che ad un singolo gene con un grande effetto. Un esempio è l'infarto miocardico, estremamente comune nella popolazione: se si osservano diverse generazioni di una famiglia in cui più individui sono affetti da infarto miocardico, ci si può domandare se la patologia sia associata ad un nuovo gene a trasmissione mendeliana o ad un comune e scarso stile di vita dei membri della famiglia;
- delezioni o inserzioni di una o più coppie di basi potrebbero essere trascurate (al contrario di sostituzioni a singolo nucleotide) (75).

5.2. *Next generation sequencing* e malattie comuni e complesse

Le malattie comuni e complesse sono per lo più la conseguenza dell'effetto di varianti genetiche rare con ampio effetto sul fenotipo malattia. La distinzione tra varianti genetiche “a bassa frequenza” (compresa tra 1:1000 e 1:20) e varianti genetiche rare ($<1:1000$), è molto importante al fine di valutare l'approccio di indagine più appropriato.

Varianti genetiche a bassa frequenza prediligono infatti l'uso dei cosiddetti “*exome chip*”. In questo caso è possibile catalogare tutte le varianti che segregano in una popolazione, con una frequenza pari o superiore a 1:1000, genotipizzando direttamente queste varianti e verificandone l'associazione con la malattia. In quanto tecnica di genotipizzazione, questa non può essere considerata alla pari di un sequenziamento vero e proprio.

Le varianti genetiche rare possono invece essere individuate solo mediante tecniche di sequenziamento genetico, e mediante il loro raggruppamento e studio in modo aggregato. Attualmente la sfida più grande è quella di usare le tecniche di sequenziamento per consentire la scoperta *de novo* di geni non noti che contribuiscono a tratti complessi. Mediante *exome sequencing* gli esomi possono essere sequenziati rapidamente e le sostituzioni a singolo nucleotide possono essere “chiamate” con precisione. I principali ostacoli sono per lo più di natura statistica: queste varianti, in quanto rare, non possono essere analizzate in modo individuale dovendo essere associate ad un determinato fenotipo, bensì devono essere aggregate a varianti rare simili, testate in modo collettivo e infine associate al fenotipo finale. Relativamente a questa metodologia sorgono tuttavia alcune domande: ad esempio se è necessario aggregare diverse varianti di uno stesso gene o varianti di geni diversi ma appartenenti alla stessa via di segnalazione; oppure se è necessario impostare lo studio classificando le varianti in funzione del tipo di mutazione (non sinonime, non senso, sinonime, di *splice*); oppure se è necessario imporre una soglia di frequenza entro la quale si definisce il termine “raro”, ad esempio $<1\%$. Infine, tenendo conto del fatto che la maggior parte delle mutazioni rare sono di tipo *missenso* (delle quali circa $1/5$ deleterie per la funzione della proteina, $1/2$ moderatamente deleterie e $1/4$ neutre), sarebbe certamente utile dal punto di vista statistico distinguere tra quelle che influenzano la funzione della proteina e quelle che non la influenzano. Questa discriminazione, in aiuto alla quale esistono diversi *software*, permette infatti di eliminare il “rumore di fondo” dell'analisi, aumentandone il potere statistico (75).

6. NUOVI GENI CORRELATI AL METABOLISMO DEL COLESTEROLO

Poiché i primi studi GWAS sui lipidi sono emersi nel 2008, è solo a partire dagli anni successivi che sono stati individuati nuovi geni associati al metabolismo del colesterolo LDL e/o alla cardiopatia ischemica. In particolare, grazie agli studi GWA è stato possibile correlare quest'ultimi ai loci *ABCG5/8* (*ATP-binding cassette G5/G8*), *NPC1L1* (*Niemann-Pick C1 like 1*), *ABO*, *MYLIP* (*Myosin regulatory light chain interacting protein*), *CYP7A1* (*Cytochrome P450 family 7 subfamily A member 1*) e *INSIG2* (*insulin induced gene 2*). Infine le più innovative tecniche di sequenziamento dell'esoma hanno permesso l'identificazione di un nuovo gene associato all'ipercolesterolemia familiare, noto come *LIPA* (*lysosomal acid lipase A*), e di un nuovo gene associato al metabolismo del colesterolo LDL, noto come *TM6SF2* (*transmembrane 6 superfamily member 2*). Ulteriori geni associati al colesterolo totale o LDL sono i geni *APOE* (*apolipoprotein E*) e *ANGPTL3* (*angiopoietin like 3*), quest'ultimo altresì identificato mediante *exome sequencing* in pazienti con ipercolesterolemia familiare (77).

6.1 I geni *ABCG5/ABCG8* e il gene *NPC1L1*

I geni *ABCG5* e *ABCG8*, localizzati nel cromosoma 2p21 e separati da sole 140 paia di basi, codificano per due trasportatori di membrana appartenenti alla superfamiglia dei trasportatori ABC (*ATP-binding cassette*). Le proteine *ABCG5/8*, una volta espresse, si organizzano in un eterodimero la cui funzione è quella di trasportare il colesterolo da enterociti ed epatociti rispettivamente nell'intestino e nella bile; *ABCG5/8* agisce quindi riducendo i livelli plasmatici di colesterolo LDL ed aumentandone la secrezione biliare. Mutazioni a carico delle proteine *ABCG5/8* sono state identificate nella sisterolemia, malattia autosomica recessiva caratterizzata da un aumentato assorbimento intestinale e una ridotta secrezione biliare degli steroli della dieta, ipercolesterolemia e malattia coronarica ad insorgenza precoce (77).

Il gene *NPC1L1* (*Niemann-Pick C1 like 1*), localizzato in posizione 7p13, svolge invece una funzione opposta ad *ABCG5/8*. *NPC1L1* media l'*up-take* del colesterolo dall'intestino e dalla bile rispettivamente in enterociti ed epatociti, determinando un aumento dei livelli plasmatici di colesterolo LDL e riducendone la secrezione biliare. Poiché la sua inattivazione permette quindi di abbassare i livelli plasmatici di colesterolo LDL e ridurre il rischio di aterosclerosi, *NPC1L1* ha da sempre suscitato un grande interesse dal punto di vista terapeutico. Degno di nota è un importante studio pubblicato nel 2014 dalla prestigiosa rivista *New England Journal*

of Medicine che, grazie al sequenziamento dell'intero esoma di NPC1L1, ha identificato mutazioni genetiche in grado di inattivare l'espressione del gene stesso ottenendo il medesimo risultato terapeutico dell'Ezetimibe, noto farmaco assunto per il trattamento dell'ipercolesterolemia che agisce inibendo selettivamente NPC1L1 (78).

Infine è stato osservato che ABCG5/8 e NPC1L1, svolgono un ruolo chiave sia nell'infarto miocardico sia nella formazione di calcoli biliari, ovvero due patologie apparentemente non correlate tra loro. Infatti, la presenza di mutazioni *gain-of-function* (GOF) in ABCG5/8 e la corrispondente presenza di mutazioni *loss-of-function* (LOF) nel gene NPC1L1 riduce il rischio di infarto miocardico ma aumenta il rischio di sviluppare calcoli biliari e patologie a carico delle vie biliari. Analogamente, mutazioni LOF a carico del gene ABCG5/8 sono causa di un aumento dei livelli plasmatici di colesterolo LDL e quindi di un aumentato rischio di infarto miocardico, sebbene si osservi un rischio ridotto di patologie a carico delle vie biliari (77).

6.2 Il gene ABO

Il gene *ABO*, localizzato sul braccio lungo del cromosoma 9 in posizione 34.2, codifica per le glicosiltransferasi del gruppo sanguigno ABO. L'esistenza di un'associazione epidemiologica tra il sistema del gruppo sanguigno ABO e i livelli plasmatici del colesterolo, i quali risultano particolarmente aumentati nei soggetti con gruppo sanguigno non-0, è stata dimostrata da quasi 50 anni; inoltre recenti studi GWA hanno avvalorato questa associazione. Coerentemente, studi epidemiologici e genetici hanno dimostrato come soggetti "non-0" presentino anche un forte rischio di sviluppare patologie ischemiche, come clinicamente dimostrato dallo studio effettuato su eritrociti legati alla apolipoproteina B (ery-ApoB). In quest'ultimo caso si è osservato che soggetti con gruppo sanguigno non-0, e bassi livelli di ery-ApoB, presentano un elevato rischio aterosclerotico. È stata quindi ipotizzata una stretta correlazione tra il sistema ABO, i livelli plasmatici il colesterolo e il rischio aterosclerotico. Inoltre, il gruppo sanguigno 0, essendo causa di un fenotipo privo di attività glicosiltransferasica, è paragonabile ad una variante LOF del gene stesso, il cui effetto si traduce in una riduzione dei livelli plasmatici del colesterolo e quindi in una protezione dal rischio aterosclerotico (77).

6.3 Il gene MYLIP

Recenti studi GWA hanno suggerito come il gene *MYLIP*, noto anche come *IDOL* (*Inducible degrader of the LDL receptor*), contribuisca alla variazione dei livelli plasmatici di colesterolo

LDL. Il gene, localizzato in posizione 6p22.3, codifica per un'E3-ubiquitina ligasi, importante regolatore post-trascrizionale, la cui espressione è mediata dall'attivazione del recettore del fegato LRX in presenza di elevati livelli intracellulari di colesterolo. Una volta espressa, MYLIP agisce ubiquitinando LDLR e inducendone la degradazione; MYLIP limita quindi l'assorbimento del colesterolo LDL mediante una via di segnalazione indipendente dal controllo trascrizione di LDLR da parte di steroli o elementi di regolazione degli steroli. Degno di nota, è un importante studio condotto da Sorrentino e collaboratori nel 2013, il quale ha identificato la prima mutazione LOF di *MYLIP* (p.Arg266X), responsabile di una perdita totale della sua attività E3-ubiquitina ligasica; questa mutazione è quindi causa di aumento dell'*up-take* del colesterolo e della sua riduzione a livello plasmatico (79).

6.4 Il gene *CYP7A1*

Questo gene, localizzato in posizione 8q12.1, codifica per l'enzima CYP7A1, il quale svolge un ruolo cruciale nel catabolismo epatico del colesterolo e nella biosintesi dell'acido biliare. Nonostante diversi GWAS abbiano dimostrato come CYP7A1 sia fortemente correlato ai livelli di colesterolo totale ed LDL, sono pochi gli studi che hanno indagato gli effetti derivanti dalla sua assenza. Alcuni studi di meta-analisi hanno confermato l'associazione tra la variante genetica -204A>C, nella regione promotrice del gene, e i livelli di colesterolo LDL in pazienti con malattia alle vie biliari. È stata altresì descritta una associazione tra alcune varianti genetiche in CYP7A1 e una riduzione del colesterolo LDL in risposta al trattamento con statine (77).

6.5 Il gene *INSIG2*

Il gene *INSIG2*, localizzato sul braccio lungo del cromosoma 2 in posizione 14.1, codifica per una proteina del reticolo endoplasmatico che inibisce il processamento delle proteine SREBP (*sterol regulatory element-binding protein*) mediante il legame con le proteine SCAP (*SREBP-cleavage-activating protein*). Sebbene già da diverso tempo fosse noto il ruolo di *INSIG2* nel metabolismo del colesterolo LDL, solo recentemente grazie a studi GWAS e alla mappatura del gene, è stato possibile dimostrare la sua stretta correlazione con i livelli plasmatici di colesterolo LDL. Dal punto di vista funzionale, infatti, *INSIG2* inibisce la trascrizione dei geni coinvolti nell'*up-take* e nella sintesi del colesterolo, mediante legame con gli ossisteroli e mediante la conseguente inibizione delle proteine SREB. La sintesi del colesterolo può essere

altresì compromessa a seguito del legame di INSIG2 con la reduttasi HMG-CoA, enzima chiave della biosintesi endogena del colesterolo, che viene successivamente ubiquitinata e degradata. Il recente sequenziamento dell'esoma del gene in pazienti affetti da ipercolesterolemia familiare, ma privi di mutazioni nei classici geni associati a questa patologia (LDLR, APOB, PCSK9), ha inoltre permesso l'identificazione di nuove varianti di *INSIG2*, avvalorando il ruolo funzionale che questa proteina svolge nel metabolismo del colesterolo LDL (77).

6.6 Il gene *LIPA*

Il gene *LIPA*, localizzato in posizione 10q23.31, codifica per la lipasi acida lisosomiale A. Dal punto di vista funzionale questo enzima agisce idrolizzando a livello lisosomiale gli ester del colesterolo e i trigliceridi. Mutazioni nel gene *LIPA* sono causa ormai nota della malattia da accumulo di ester del colesterolo (CESD), rara patologia autosomica recessiva caratterizzata da un progressivo accumulo di ester del colesterolo e trigliceridi nel fegato, milza e altri organi, dislipidemia e aterosclerosi precoce. Sebbene il gene *LIPA* sia stato identificato mediante sequenziamento dell'esoma in soggetti affetti da ipercolesterolemia familiare, e diversi GWAS abbiano correlato questo gene alla sindrome coronarica, deve essere fatto ancora molto per comprendere il suo ruolo nel metabolismo del colesterolo (77).

6.7 Il gene *TM6SF2*

Il gene *TM6SF2*, localizzato in posizione 19q13.11, è stato individuato nel 2014, ovvero in corrispondenza dei primi esperimenti di sequenziamento dell'esoma effettuati con lo scopo di individuare nuovi geni associati al metabolismo del colesterolo. Questo gene non solo è correlato ai livelli plasmatici di colesterolo LDL ma anche al rischio di infarto miocardico. In primo luogo è stato dimostrato che l'*over-espressione* di *TM6SF2* nel modello murino determina un incremento dei livelli plasmatici del colesterolo totale, del colesterolo LDL e dei trigliceridi, così come una riduzione dei livelli plasmatici del colesterolo HDL. Parallelamente è stata dimostrata la correlazione tra varianti del gene *TM6SF2* e il rischio di steatosi epatica non alcolica (NAFLD); dal punto di vista funzionale, infatti, la proteina *TM6SF2* aumenta la secrezione del colesterolo VLDL dagli epatociti e riduce il contenuto di gocce lipidiche a livello epatico. *TM6SF2* svolge quindi un importante ruolo sia nella malattia coronarica sia nella steatosi epatica non alcolica: soggetti con varianti LOF del gene si caratterizzano infatti

da ridotti livelli plasmatici di colesterolo totale e trigliceridi, e una corrispondente riduzione del rischio cardiovascolare, ma allo stesso tempo sono maggiormente suscettibili a NAFLD (77).

6.8 APOE

Il gene *APOE*, localizzato in posizione 19q13.32, codifica per una glicoproteina di 34 kDa, nota componente di chilomicroni e delle lipoproteine VLDL, IDL, LDL e HDL. APOE è coinvolta nel riconoscimento di tutti i componenti della famiglia di recettori dell'LDL, contribuendo così in modo significativo alla variazione del colesterolo plasmatico. Degna di nota è la mutazione p.Leu167del, riconosciuta essere una delle possibili cause, sebbene rara, dell'ipercolesterolemia familiare. Dal punto di vista funzionale questa mutazione causa una destabilizzazione del dominio di legame ad alfa-elica di APOE, una ridotta espressione della sua forma funzionale e la conseguente riduzione del catabolismo del colesterolo. La delezione p.leu167 in *APOE* è stata altresì associata all'iperlipidemia combinata familiare e l'istiocitosi blu-mare; ciononostante la penetranza con la quale si manifesta questa variante genetica è strettamente dipendente dalla presenza di altri geni modificatori (16).

6.9 Il gene ANGPTL3

Il gene *ANGPTL3*, localizzato sul braccio corto del cromosoma 1 in posizione 31.3, codifica per una importante proteina principalmente espressa nel fegato e strutturalmente molto simile alle angiopoietine. A differenza di quest'ultime, tuttavia, non si lega con alta affinità ai recettori delle angiopoietine presenti sulla membrana delle cellule endoteliali, e hanno scarsa o nulla attività nell'induzione dell'angiogenesi. In condizioni fisiologiche ANGPTL3 inibisce la lipasi lipoproteica e la lipasi endoteliale, che hanno rispettivamente la funzione di idrolizzare i trigliceridi contenuti in chilomicroni e VLDL, e di idrolizzare i fosfolipidi delle HDL; in altre parole ANGPTL3 agisce principalmente aumentando i livelli plasmatici di trigliceridi e colesterolo. Studi condotti da Musunuru e collaboratori in soggetti affetti da ipolipidemia combinata familiare hanno individuato due rare mutazioni non senso nel gene *ANGPTL3*, S17X e E129X, responsabili di una sostanziale riduzione dei livelli plasmatici di colesterolo; difatti la mancata inibizione delle lipasi lipoproteiche e delle lipasi endoteliali, che si verifica in presenza di mutazioni LOF in *ANGPTL3*, si traduce in un aumento dell'attività di questi due enzimi e una conseguente riduzione dei livelli di colesterolo plasmatico. ANGPTL3 rappresenta dunque

un interessante bersaglio terapeutico per l'ipercolesterolemia e per proteggere verso la malattia cardiovascolare (76).

7. IL CONCETTO DI *MISSING HERITABILITY*

La cardiopatia ischemica è definita una malattia complessa ad eziologia multifattoriale, ovvero influenzata sia da fattori ambientali sia genetici. Nonostante i fattori di rischio ambientali siano noti da diverso tempo, solo le recenti innovazioni scientifiche e tecnologiche in campo genetico hanno consentito di identificare varianti genetiche associate in maniera significativa alla cardiopatia ischemica, rinforzando il ruolo che i fattori genetici hanno nello sviluppo e progressione della stessa. Tuttavia, sebbene l'ereditarietà stimata per questa malattia sia del 40-60%, la caratterizzazione delle varianti genetiche fino ad oggi individuate ha permesso di comprenderne solo il 15-20% (80, 81). Questa disparità osservata tra ereditarietà attesa e realmente osservata è nota con il termine "*missing heritability*": fenomeno che si manifesta con proporzioni pressoché paragonabili per quasi tutte le malattie il cui rischio genetico è stato identificato mediante studi GWAS (82). Quest'ultimi rappresentano infatti un potente strumento di analisi genetica per le malattie comuni e complesse, ma mostrano anche considerevoli limiti, alcuni dei quali sono rappresentati da un'imprecisa definizione del fenotipo da studiare o da una discutibile selezione del gruppo di controllo (83, 84). Non solo, ma gli *array* di cui si avvalgono gli studi GWAS per lo studio del genotipo non sono in grado di osservare alcune varianti strutturali causative della malattia, così come non permettono di individuare e caratterizzare le interazioni tra multipli geni. Infine, è importante osservare che alla base degli studi di GWAS vi è il concetto "*common disease, common variant*", secondo il quale le malattie comuni sono conseguenti a varianti alleliche comuni, ovvero presenti almeno nel 5% della popolazione (85-87). Una porzione consistente del fenomeno dell'ereditarietà mancante sembra proprio essere conseguente a tutte quelle varianti genetiche con una frequenza bassa, ovvero compresa tra lo 0,5% e il 5%, o addirittura rare, inferiore allo 0,5%. Queste varianti, a causa della loro frequenza, sfuggono alle analisi GWAS; inoltre non sono responsabili di un effetto tale da poter essere individuate mediante analisi di *linkage* in studi di famiglia (Figura 4) (87, 88). Precisamente è importante osservare che nel momento in cui la frequenza allelica di una variante è inferiore allo 0,5%, la rilevazione di una sua associazione

con la malattia diventa pressoché improbabile, a meno che non sia causa di un ampio effetto sul fenotipo come nel caso dei disordini monogenici (82).

Le varianti a bassa frequenza e rare sono associate ad un rischio di sviluppare la malattia raddoppiato, o addirittura triplicato, rispetto ad un soggetto non portatore; si tratta di mutazioni che sebbene non seguano una segregazione mendeliana, potrebbero spiegare una quota consistente di quella ereditarietà ad oggi ancora ignota (89)

È essenziale comprendere interamente l'entità del contributo della genetica alle malattie complesse, quali le malattie cardiovascolari, poiché la maggior parte delle differenze individuali legate alla suscettibilità di una determinata malattia sono provocate dalle caratteristiche del patrimonio genetico di un individuo, e la comprensione di questa variabilità potrà contribuire ad una migliore prevenzione, diagnosi e trattamento farmacologico della malattia (82). In conclusione, sebbene gli studi GWAS si siano dimostrati un efficiente sistema per ricercare l'ereditarietà delle patologie complesse, solo il sequenziamento diretto di quelle regioni che hanno mostrato una forte associazione con le varianti comuni, o il sequenziamento diretto di soggetti con fenotipo estremo, potranno permettere la caratterizzazione delle varianti realmente causative della malattia e aiutare a comprendere definitivamente l'ereditarietà che ne sta alla base (75). Le tecniche di sequenziamento di ultima generazione, diversamente dal sequenziamento Sanger, consentono l'analisi di milioni di sequenze in parallelo, incrementando notevolmente la velocità di analisi e la mole dei dati generati (90).

SCOPO DELLA TESI

Tenuto conto dello stato dell'arte appena illustrato, lo scopo di questo lavoro è stato quello di valutare l'associazione di varianti genetiche rare, che predispongono a ipercolesterolemia o ipocolesterolemia, con il rischio di infarto miocardico acuto e il fenotipo intermedio della malattia, ovvero i livelli plasmatici di colesterolo LDL e il *burden* aterosclerotico coronarico. Al fine di raggiungere questo obiettivo, è stata analizzata mediante *exome sequencing* la presenza di varianti genetiche con una frequenza $< 1\%$ in una popolazione estremamente preziosa dal punto di vista genetico: 4016 soggetti, di cui 2008 individui sani e 2008 pazienti con infarto miocardico giovanile, nei quali è probabile che l'eziopatogenesi della malattia abbia prevalentemente una componente genetica (Figura 5). Infatti, minore è l'età alla quale si manifesta la cardiopatia ischemica, maggiore è la possibile influenza genetica sulla patologia e, di conseguenza minore sarà l'influenza dei fattori ambientali.

MATERIALI E METODI

1. POPOLAZIONE DELLO STUDIO E RACCOLTA DEI CAMPIONI

La popolazione dello studio è costituita da 2008 casi e 2008 controlli arruolati tra il 1988 e il 2002 in 125 Cardiologie su tutto il territorio nazionale, nell'ambito dell' "*Italian Genetic Study of Early-onset Myocardial Infarction*". Sono stati definiti "casi" pazienti che hanno avuto un infarto miocardico giovanile, ovvero entro i 45 anni di età; sono stati definiti "controlli" soggetti sani selezionati affinché fossero paragonabili ai casi per età, sesso e origine geografica. In tutti i soggetti, dopo aver ottenuto il consenso informato, è stato eseguito un prelievo di sangue per l'analisi del DNA. Le variabili cliniche di interesse (familiarità per malattie cardiovascolari, fattori di rischio cardiovascolari, stile di vita, terapia farmacologica) sono state raccolte in una scheda raccolta dati cartacea e informatizzata; il consenso informato è stato ottenuto rispettando i requisiti del comitato etico locale, della Società Italiana di Genetica Umana (SIGU) e della Società Europea di Genetica Umana (ESHG).

I prelievi di sangue sono stati eseguiti via venosa periferica, raccolti in tre provette contenenti sodio citrato (0.106 M) e sottoposti a centrifugazione per separare il plasma dai globuli rossi. Il plasma ottenuto è stato diviso in cinque aliquote e conservato a -80°C. Il DNA è stato estratto dai globuli bianchi mediante metodo "*salting out*" ed è stato conservato in alcool a -80°C.

2. ANALISI MOLECOLARE

I campioni selezionati sono stati successivamente sottoposti ad analisi molecolare *whole exome sequencing* (WES); nel caso del gene *PCSK9*, poiché l'esone 1 non è stato efficacemente sequenziato, si è proceduto attraverso analisi della variante R46L (1:55505647_G/T) mediante *Real Time PCR*.

L'analisi molecolare WES è stata eseguita presso il Broad Institute (MIT, Boston). I DNA estratti sono stati sottoposti ad analisi quantitativa mediante *Picogreen assay* (Invitrogen) e conservati a -20°C. Sono stati esclusi tutti i campioni i cui valori preliminari di massa totale, concentrazione, integrità o qualità del DNA erano troppo bassi. La libreria genica e la successiva selezione ibrida in soluzione sono stati realizzati a partire da 3 µg di DNA genomico. Le librerie ottenute sono state quantificate utilizzando qPCR (KAPA Biosystems), normalizzate a 2nM e successivamente sottoposte a denaturazione con 0,1 N di NaOH. L'amplificazione dei *cluster* è stata effettuata

attendendosi al protocollo Illumina; ad ognuna delle celle a flusso è stato quindi addizionato il fluoroforo SYBR *green*. Le celle sono state successivamente visualizzate a microscopio ottico al fine di confermare la densità dei *cluster* e sottoposte a sequenziamento Genome Analyzer II o HiSeq 2.000 (76 cicli).

3. STUDIO E SELEZIONE DEI GENI

Sono stati presi in considerazione esclusivamente i geni che soddisfacevano i seguenti criteri:

- Geni classicamente associati all'ipercolesterolemia familiare (16, 76, 77)
- Geni associati a ipercolesterolemia secondo le informazioni fornite dal *database* OMIM (*Online Mendelian Inheritance in Man*) (91)
- Geni associati al metabolismo del colesterolo o colesterolo LDL (77)

Questa prima selezione ci ha permesso di individuare 20 geni.

Successivamente è stata effettuata una seconda selezione, ovvero sono stati selezionati solo quei geni per i quali fosse chiaro l'effetto delle mutazioni *loss-of-function* (LOF), ovvero distruttive per la funzione della proteina, sui livelli del colesterolo. Tale analisi *in silico* ci ha permesso di restringere la nostra attenzione su 15 geni: *ABCG5*, *ABCG8*, *APOB*, *APOE*, *CYP7A1*, *INSIG2*, *LDLR*, *LDLRAP1*, *LIPA*, *ABO*, *ANGPTL3*, *MYLIP*, *PCSK9*, *TM6SF2*, *NPC1L1*. In tabella 2 sono stati riportati i dettagli relativi alla posizione cromosomica dei geni selezionati, alla loro associazione con il colesterolo totale o LDL e la cardiopatia ischemica, alla loro identificazione mediante tecnica *exome sequencing* in pazienti con ipercolesterolemia familiare, all'esistenza di studi a supporto di un loro ruolo nel metabolismo del colesterolo LDL e il loro effetto sui livelli del colesterolo LDL quando sottoposti a mutazioni LOF.

4. STUDIO DELLE VARIANTI RARE

In seguito alla selezione dei geni sono state prese in considerazione le sole varianti rare, ovvero le varianti con un'incidenza nella popolazione inferiore all'1%. Le varianti *missenso* sono state analizzate avvalendosi di algoritmi di predizione di patogenicità, disponibili all'interno del *database* dbNSFP (92). In particolare sono stati utilizzati i seguenti algoritmi: SIFT (93), *PolyPhen2* (94), *MutationTaster* (95), *MutationAssessor* (96), LTR (*likelihood ratio test*) (97), FATHMM (*Functional Analysis Through Hidden Markov Model*) (98). Tale analisi ci ha permesso di predire la

patogenicità delle varianti genetiche considerate. Infine, le mutazioni *missense* predette come deleterie dagli algoritmi sono state ulteriormente indagate avvalendosi di *database* di pubblico accesso, quali LOVD (*Leiden Open-source Variation Database*) (99), ExPASY (*Expert Protein Analysis System*) (100), ClinVar (101) e UMD (*Universal Mutation Database*) (102). Queste varianti sono state considerate varianti distruttive LOF solo se era già descritto un loro ruolo nel metabolismo dell'LDL da dati ottenuti da studi di famiglia o popolazione.

5. DETERMINAZIONE DEI LIVELLI PLASMATICI DI COLESTEROLO E VALUTAZIONE DELLE STENOSI CORONARICHE

Il colesterolo totale, il colesterolo HDL e i livelli dei trigliceridi sono stati misurati da campioni di sangue venoso in condizioni di digiuno secondo le procedure standard. I livelli di colesterolo plasmatico LDL sono stati calcolati con la formula di Friedewald; il valore dei trigliceridi è stato assunto come mancante se superiore a 400 mg/dl (4,5 mmol per litro). Tutte le terapie volte ad una riduzione dei livelli del colesterolo sono state registrate.

Sono state definite normali coronarie prive di qualsiasi restrizione del diametro. Sono state considerate stenosi non significative riduzioni del diametro coronarico <70% (50% nel caso del tronco comune); sono state considerate stenosi significative restrizioni $\geq 70\%$ (50% o più nel caso del tronco comune). È stata definita malattia monovasale, la presenza di stenosi significativa in un solo vaso coronarico; è stata definita malattia multivasale la presenza di stenosi significativa in almeno due delle tre arterie principali.

Il *burden* aterosclerotico coronarico è stato misurato avvalendosi del *Duke Coronary Artery Disease Index* (103), che è stato calcolato in tutti i pazienti sulla base dell'angiografia coronarica effettuata in corso di ospedalizzazione degli stessi per l'evento indice (infarto miocardico)

6. ANALISI STATISTICA

Le variabili continue sono state riportate in termini di media e deviazione standard; le variabili dicotomiche o categoriche sono state riportate in termini di frequenza assoluta e relativa.

Sono stati calcolati gli Odds Ratio (OR) e gli intervalli di confidenza (CI) relativi all'associazione della variante genetica con l'ipercolesterolemia o ipocolesterolemia plasmatica e al rischio di avere un infarto miocardico; la significatività dei valori di OR è stata valutata attraverso un test statistico (χ^2 -test). Le analisi sono state condotte utilizzando R-Scripts versione di R-3.4.1.

RISULTATI

1. SEQUENZIAMENTO E SELEZIONE DEI GENI

Complessivamente è stato possibile ottenere l'esoma per 3532 soggetti (1790 casi e 1742 controlli). Infatti, rispetto alla coorte iniziale di 4016, per 484 campioni non è stato possibile portare a termine il sequenziamento per cause tecniche legate alla procedura di sequenziamento o scarsa qualità del DNA estratto.

I risultati relativi all'analisi delle variabili rare nei 15 geni selezionati sono stati riportati nella Tabella 3. In totale sono state identificate 47 varianti distruttive in 82 soggetti e 53 varianti *missenso* in 83 soggetti (valutati globalmente tra casi e controlli).

Allo scopo di chiarire l'effetto funzionale di mutazioni *missenso* predette come deleterie dalle analisi bioinformatiche effettuate, abbiamo voluto verificare se vi fossero ulteriori evidenze relativamente ad un loro possibile ruolo patogenetico nei *database* di pubblico accesso. Le varianti precedentemente note svolgere un ruolo nel metabolismo dell'LDL che, sulla base di dati ottenuti su studi di famiglia o popolazione, risultavano avere un effetto distruttivo sulla funzione della proteina, sono state successivamente considerate mutazioni LOF (Tabella 3)

È importante osservare che quasi tutti i soggetti sono portatori di almeno una variante in studio nello stato eterozigote. Le sole eccezioni sono rappresentate da un caso omozigote per una mutazione LOF in *LDLR*, un soggetto portatore di due varianti nello stesso gene (*NPC1L1*), e quattro soggetti portatori di due mutazioni eterozigoti in geni differenti (Tabella 4).

2. EFFETTO DELLE VARIANTI GENETICHE SUL RISCHIO DI INFARTO MIOCARDICO

Al fine di verificare se le varianti che predispongono ad ipercolesterolemia sono più frequenti in chi ha avuto un infarto miocardico giovanile rispetto ad un soggetto sano e di pari sesso, età ed origine geografica, siamo andati ad osservare la distribuzione di tali varianti genetiche nel campione dei casi e nel campione dei controlli. Il risultato ottenuto mostra che le varianti genetiche predisponenti verso l'ipercolesterolemia sono più frequenti nei casi rispetto ai controlli. In particolare 76 casi e 26 controlli sono portatori di almeno una variante deleteria o *missenso* associata ad un aumento dei livelli di colesterolo LDL (Tabella 5): tali varianti risultano associate pertanto ad una maggiore probabilità di andare incontro ad un infarto miocardico giovanile (OR = 2.93, 95% CI 1,89-4,67;

$p < 0,000002$). Al contrario le varianti genetiche associate a ipocolesterolemia sono più frequenti nei controlli rispetto ai casi di pari età, sesso e origine geografica: infatti 57 casi e 81 controlli sono portatori di almeno una variante genetica deleteria o *missenso* che determina ipocolesterolemia (Tabella 5). Tali varianti sono associate ad una ridotta probabilità di sviluppare un infarto miocardico giovanile (OR = 0,67, 95% CI 0,47-0,91; $p < 0,025$) (Figura 6).

3. EFFETTO DELLE VARIANTI GENETICHE SUL FENOTIPO INTERMEDIO

Si è proceduto quindi ad indagare se le varianti genetiche che predispongono ad ipercolesterolemia sono di fatto associate a dei livelli più alti di colesterolo plasmatico; a tale scopo si è proceduto alla misurazione dei livelli plasmatici di colesterolo LDL per l'intera popolazione. In seguito siamo andati ad osservare quale fosse la distribuzione delle varianti predisponenti e protettive in funzione dei livelli plasmatici di colesterolo LDL nella popolazione in studio. Quello che abbiamo potuto osservare è che le varianti genetiche predisponenti sono maggiormente rappresentate in soggetti con alti livelli di colesterolo plasmatico, mentre quelle protettive sono maggiormente presenti in soggetti con livelli di colesterolo plasmatico normale. Questa condizione è illustrata dal *box-plot* mostrato in Figura 7, dove si osserva che soggetti con varianti genetiche predisponenti hanno un valore medio di colesterolo plasmatico LDL pari a 203,48 mg/dl (SD = 91,29 mg/dl), ovvero circa 1,5 volte superiore rispetto al valore medio di colesterolo plasmatico LDL dei soggetti che possiedono varianti genetiche protettive, i quali hanno un valore medio di colesterolo LDL pari a 131,02 mg/dl (SD = 41,64 mg/dl) (Figura 7).

Al fine di valutare l'effetto delle varianti genetiche sul *burden* aterosclerotico si è proseguito con il calcolo del *Duke Coronary Index*. Il valore mediano del *Duke Coronary Index* coronarico è risultato essere 37 nei 76 casi portatori di varianti genetiche associati a ipercolesterolemia, e 32 nei 57 casi portatori di varianti genetiche associate a ipocolesterolemia; infine è stato calcolato un *Duke Coronary Index* pari a 32 relativamente anche ai 1714 casi privi di varianti genetiche associate a condizioni di ipercolesterolemia o ipocolesterolemia. Tali differenze sono risultate statisticamente significative ($p < 0,002$).

In tabella 6 sono stati riportati tutti i dettagli relativi alle variabili cliniche e al grado di stenosi coronarica relativi ai soggetti portatori di varianti associate a ipercolesterolemia e ipocolesterolemia nei geni selezionati.

DISCUSSIONE

Questo studio è stato svolto al fine di valutare l'associazione tra varianti genetiche rare in 15 geni, noti influenzare i livelli di colesterolo LDL plasmatico, e il rischio di infarto miocardico. L'obiettivo è stato raggiunto avvalendosi di un eccezionale modello sperimentale dal punto di vista genetico, in quanto costituito da una popolazione di 2008 soggetti sani e 2008 pazienti con infarto miocardico giovanile, ovvero pazienti nei quali si presuppone, data l'età di insorgenza dell'infarto, un più importante contributo all'evento indice dei fattori genetici rispetto ai fattori ambientali.

I risultati dello studio confermano l'ipotesi che le varianti genetiche, che predispongono a ipercolesterolemia o ipocolesterolemia, si associno sia al rischio di infarto miocardico sia al fenotipo intermedio, definito come colesterolo LDL plasmatico e il *burden* aterosclerotico coronarico valutato mediante *Duke Coronary Score*, sia in senso protettivo sia in senso predisponente.

L'infarto miocardico è una malattia multifattoriale, la cui eziopatogenesi deriva dall'interazione tra fattori di rischio ambientali e fattori di rischio genetici. I fattori ereditari possono contribuire fino al 30-60% della variazione del rischio, e studi eseguiti su molte popolazioni hanno identificato numerose varianti genetiche associate significativamente alle malattie cardiovascolari (78, 80, 81, 104-107). L'associazione tra varianti genetiche e prognosi di cardiopatia ischemica è stata ben definita nei pazienti con ipercolesterolemia familiare (FH), un disturbo genetico caratterizzato da un'errata capacità di rimozione dal sangue del colesterolo LDL e, nella maggior parte dei casi, ereditato in maniera autosomica dominante conseguentemente all'insorgenza di mutazioni nei geni *LDLR*, *APOB* e *PCSK9*; la forma recessiva di questa malattia è invece provocata da mutazioni nel gene *LDLRAP1* (16). È stato osservato che circa 1 individuo su 500 è affetto da FH e, se non trattati dal punto di vista terapeutico, hanno una maggiore probabilità di sviluppare precocemente malattie cardiovascolari. Inoltre, diversi studi hanno indicato che circa il 5% dei pazienti che hanno avuto un infarto miocardico prima dei 60 anni di età sono affetti dalla forma eterozigote di FH; mentre quasi il 50% dei pazienti con FH eterozigote, e non trattati farmacologicamente, vanno incontro a infarto miocardico entro i 60 anni di vita. Queste considerazioni suggeriscono che elevati livelli di colesterolo LDL rappresentano condizione sufficiente allo sviluppo della cardiopatia ischemica, e supportano l'ipotesi che può essere importante distinguere l'ipercolesterolemia geneticamente determinata da fenotipi clinicamente simili, al fine di garantire un'adeguata gestione terapeutica e un'opportuna consulenza genetica. Infatti, i fattori di rischio genetici sono presenti sin dalla nascita in un individuo, ragione per la quale questi possono avere un maggiore impatto sul rischio di infarto

miocardico giovanile ed eventi ischemici; in queste condizioni si osserva un'esposizione cumulativa nel tempo di elevati livelli plasmatici di colesterolo LDL.

Gli studi GWAS hanno focalizzato la propria attenzione sulle varianti genetiche comuni, permettendo così l'individuazione di 157 loci associati ai livelli plasmatici dei lipidi, incluso il colesterolo LDL, identificato tra i più importanti fattori di rischio delle malattie cardiovascolari. Ciononostante solo il 10-20% della variabilità totale del colesterolo LDL è attribuibile a queste varianti genetiche comuni, nessuna delle quali è stata associata ad eventi successivi (80-82). Questo progetto di ricerca ha voluto avvalersi di nuove e più innovative tecniche di sequenziamento genetico (*exome sequencing*) al fine di verificare l'influenza di varianti genetiche rare note influenzare i livelli di colesterolo plasmatico LDL sul rischio di sviluppare infarto miocardico.

Le varianti genetiche predisponenti a ipocolesterolemia e ipercolesterolemia erano presenti nel 7,4% della popolazione in studio. Le varianti genetiche associate ad un aumento dei livelli plasmatici di colesterolo LDL sono risultate associate ad un maggior rischio di infarto miocardico, le varianti genetiche associate ad una riduzione dei livelli plasmatici di colesterolo LDL sono risultate associate ad un rischio minore di infarto miocardico.

Inoltre, i livelli di colesterolo plasmatico LDL sono risultati essere significativamente più elevati nei pazienti con varianti associate ad ipercolesterolemia, e significativamente inferiori nei pazienti con varianti associate a ipocolesterolemia. Infine, pazienti con varianti genetiche associate a ipercolesterolemia presentano un *Duke Coronary Index* elevato, indicando una maggiore estensione di aterosclerosi coronarica; al contrario pazienti con varianti genetiche associate a ipocolesterolemia sono caratterizzati da un *Duke Coronary Index* inferiore.

Il risultato più significativo di questo studio è che l'analisi delle varianti genetiche, oltre ai livelli del colesterolo, permette di identificare un gruppo piccolo, ma non trascurabile, di soggetti con un rischio maggiore in una popolazione a rischio. L'importanza di identificare una popolazione ad alto rischio attraverso studi genetici è stata recentemente illustrata dal gruppo di Mega e collaboratori, che avvalendosi di un *risk score* basato su 27 varianti genetiche comuni, identificate mediante analisi GWAS, è stato in grado di individuare una sottopopolazione con un aumentato rischio di eventi coronarici che potrà beneficiare in modo particolare della terapia con statine (108).

Questo studio ha considerato esclusivamente varianti genetiche rare in geni che sono già noti in letteratura avere un effetto sui livelli di colesterolo LDL valutati in una popolazione che ha avuto un infarto miocardico giovanile. La presenza di almeno una variante associata a ipercolesterolemia è stata altresì associata ad un *Duke Index* superiore e un maggior rischio di infarto giovanile. La presenza di almeno una variante associata a ipocolesterolemia è stata altresì associata a ridotti valori del *Duke Coronary Index* e una ridotta probabilità di incorrere nell'infarto giovanile. La maggior

parte dei pazienti, privi di varianti geneticamente interessanti, presentano livelli intermedi di colesterolo LDL e una probabilità intermedia di incorrere in un infarto giovanile. Questi dati confermano che un aumentato rischio tra portatori di queste mutazioni è almeno parzialmente spiegato da una prolungata esposizione nel corso della vita ai livelli di colesterolo LDL. Questa popolazione potrà quindi beneficiare di strategie preventive più intense e trattamenti farmacologici personalizzati.

FIGURE E TABELLE

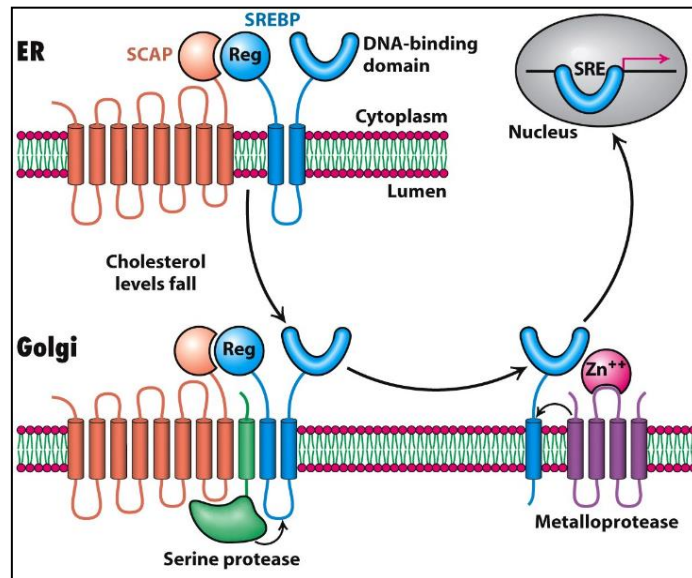


Figura 1. La via di segnalazione della proteina SREBP (14)

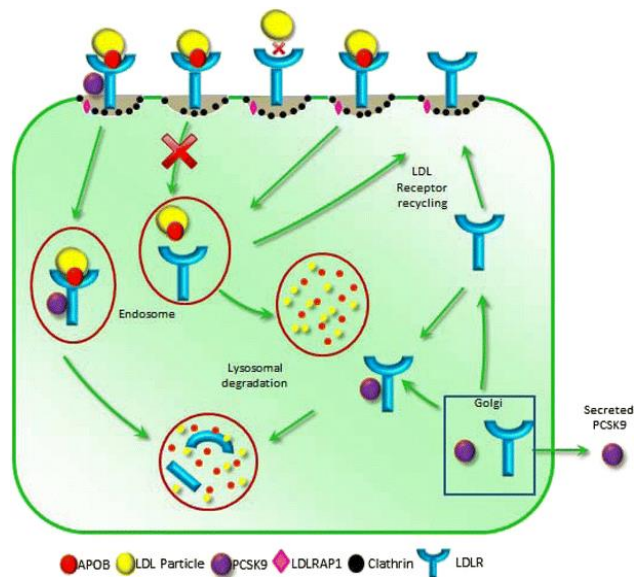


Figura 2. Internalizzazione del colesterolo LDL ed effetti nella via di segnalazione dell'LDLR a seguito di mutazioni a carico dei geni codificanti per le proteine LDLR, APOB, LDLRAP1 e PCSK9 (16)

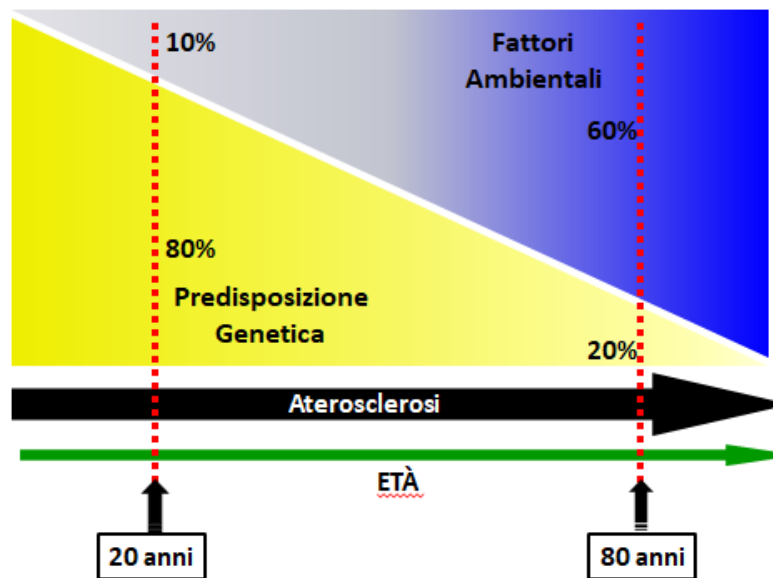


Figura 3. Multifattorialità della cardiopatia ischemica.

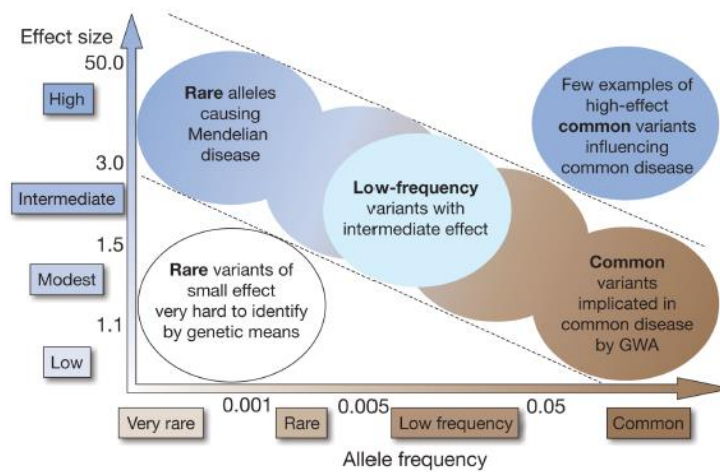


Figura 4. Capacità di identificazione delle varianti genetiche di diversa frequenza in funzione dell'effetto sul rischio di sviluppare la malattia

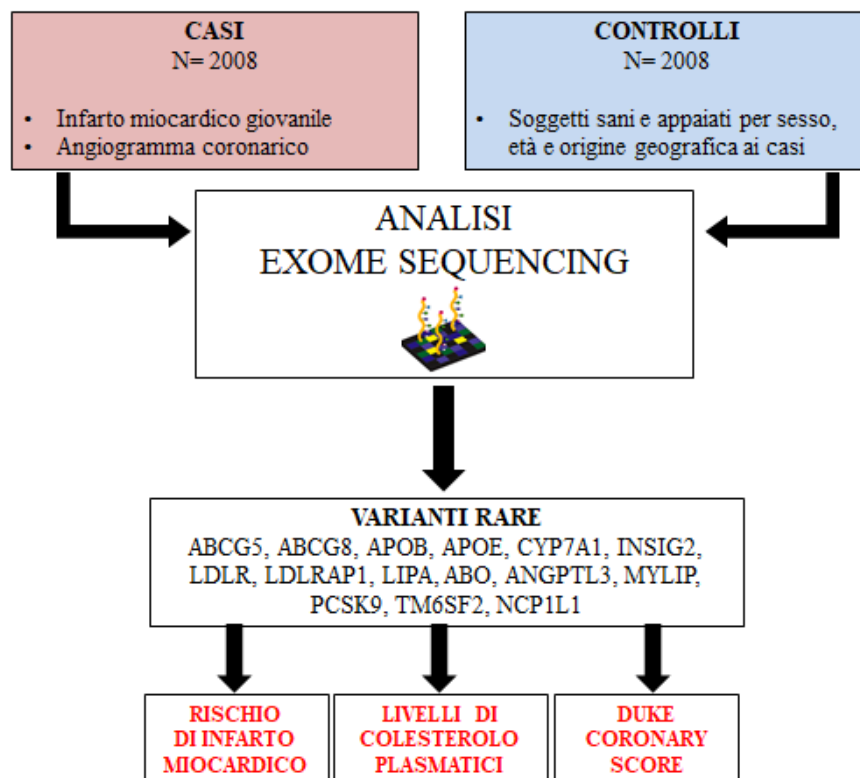


Figura 5. Illustrazione schematica dello studio

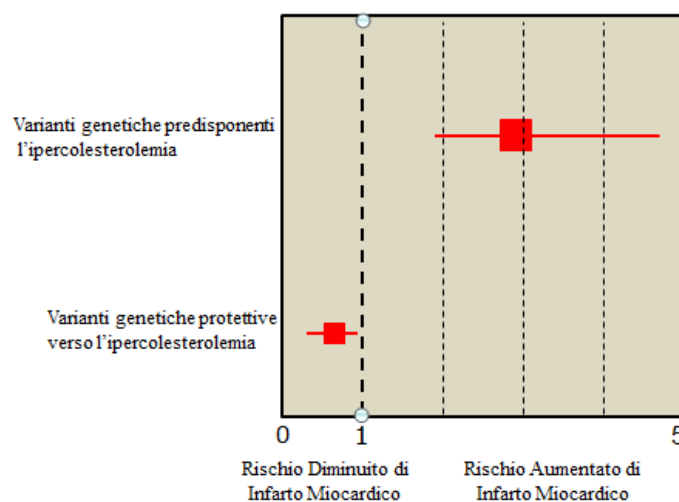


Figura 6. Rappresentazione grafica del rischio di sviluppare l'infarto miocardico giovanile sulla base della presenza delle variabili genetiche associate a ipercolesterolemia e ipocolesterolemia

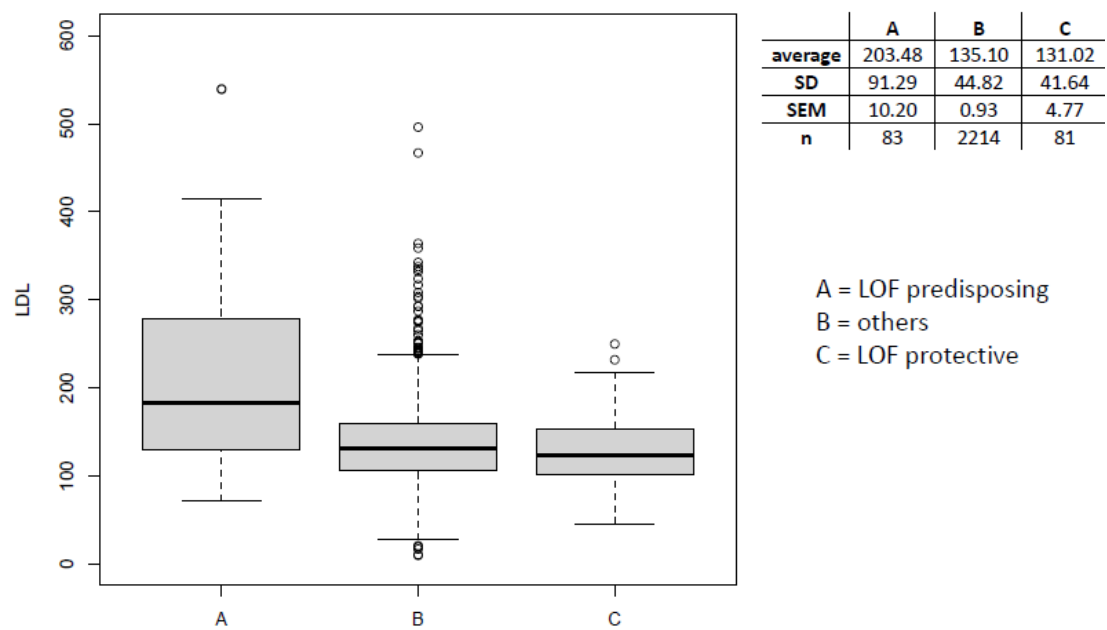


Figura 7. *Box-plot* relativo ai valori di colesterolo LDL nei gruppi A (soggetti con varianti LOF predisponenti), C (soggetti con varianti LOF protettive) B (soggetti privi delle varianti nei geni selezionati)

Tabella 1. Proprietà delle lipoproteine plasmatiche (16)

Lipoproteine plasmatiche	Densità (g ml⁻¹)	Diametro (nm)	Apolipoproteina	Ruolo fisiologico
Chilomicroni	< 0,95	75-1200	B48, C, E	Trasporto dei grassi della dieta
Lipoproteine a densità molto bassa (VLDL)	0,95 – 1,006	30 – 80	B100, C, E	Trasporto dei grassi endogeni
Lipoproteine a densità intermedia (IDL)	1,006 – 1,019	15 – 35	B100, E	Precursore delle LDL
Lipoproteina a bassa densità (LDL)	1,019 – 1,063	18 – 25	B100	Trasporto del colesterolo
Lipoproteina ad alta densità (HDL)	1,063 – 1,21	7,5 – 20	A	Trasporto invertito del colesterolo

Tabella 2 Geni selezionati per l'analisi delle varianti

Gene	Posizione (hg19)	Associazione con colesterolo totale o LDLC	Associazione con CAD	Identificazione mediante <i>exome sequencing</i> in pazienti con FH	Studi a supporto del ruolo del gene nel metabolismo dell'LDL	Effetto di mutazioni LOF sui livelli di LDL-C	Referenze
<i>ABCG5</i>	chr2:44039611-44065958	X	X	X	X	↑	1,2,3,4
<i>ABCG8</i>	chr2:44066103-44105605	X	X	X	X	↑	1,2,3,4
<i>APOB</i>	chr2:21224301-21266945	X	X	X	X	↑/↓	5,6,7,8
<i>APOE</i>	chr19:45409006-45412652	X	X	X	X	↑	6,8,9
<i>CYP7A1</i>	chr8:59402737-59412720	X	-	-	X	↑	6,8
<i>INSIG2</i>	chr2:118846050-118867597	X	-	X	X	↑	6,10,11
<i>LDLR</i>	chr19:11200038-11244505	X	X	X	X	↑	6,8,12
<i>LDLRAP1</i>	chr1:25870076-25895377	X	X	X	X	↑	6,13,14
<i>LIPA</i>	chr10:90973326-91011796	-	X	X	X	↑	8,15,16
<i>ABO</i>	chr9:136130563-136150630	X	X	-	-	↓	6
<i>ANGPTL3</i>	chr1:63063158-63071976	X	X	-	-	↓	6,17,18
<i>MYLIP</i>	chr6:16129317-16148478	X	-	-	X	↓	6,19
<i>NPC1L1</i>	chr7:44552135-44580914	X	X	-	X	↓	6, 20,22
<i>PCSK9</i>	chr1:55505149-55530526	X	X	X	X	↓	6,7,8
<i>TM6SF2</i>	chr19:19375174-19384074	X	-	-	X	↓	6,21

CAD= *cardiovascular disease*

1. Ma L, Yang J, Runesha HB, Tanaka T, Ferrucci L, Bandinelli S, Da Y. Genome-wide association analysis of total cholesterol and high-density lipoprotein cholesterol levels using the Framingham heart study data. *BMC Med Genet.* 2010 Apr 6;11:55. doi: 10.1186/1471-2350-11-55. PubMed PMID: 20370913; PubMed Central PMCID: PMC2867786

2. IBC 50K CAD Consortium. Large-scale gene-centric analysis identifies novel variants for coronary artery disease. *PLoS Genet.* 2011 Sep;7(9):e1002260. doi: 10.1371/journal.pgen.1002260. Epub 2011 Sep 22. Erratum in: *PLoS Genet.* 2-12 Aug;8(8). doi: 10.1371/annotation/120649cf-8c28-43c9-a688-c7cd65eb1aec. Roosendaal, Frits R [corrected to Rosendaal, Frits R]. PubMed PMID: 21966275; PubMed Central PMCID: PMC3178591.

3. Rios J, Stein E, Shendure J, Hobbs HH, Cohen JC. Identification by whole-genome resequencing of gene defect responsible for severe hypercholesterolemia. *Hum Mol Genet.* 2010 Nov 15;19(22):4313-8. doi:10.1093/hmg/ddq352. Epub 2010 Aug 18. PubMed PMID: 20719861; PubMed Central PMCID: PMC2957323.

4. Yu XH, Qian K, Jiang N, Zheng XL, Cayabyab FS, Tang CK. ABCG5/ABCG8 in cholesterol excretion and atherosclerosis. *Clin Chim Acta.* 2014 Jan 20;428:82-8. doi: 10.1016/j.cca.2013.11.010. Epub 2013 Nov 16. Review. PubMed PMID: 24252657.

5. Mäkelä KM, Seppälä I, Hernesniemi JA, Lyytikäinen LP, Oksala N, Kleber ME, Scharnagl H, Grammer TB, Baumert J, Thorand B, Jula A, Hutri-Kähönen N, Juonala M, Laitinen T, Laaksonen R, Karhunen PJ, Nikus KC, Nieminen T, Laurikka J, Kuukasjärvi P, Tarkka M, Viik J, Klopp N, Illig T, Kettunen J, Ahotupa M, Viikari JS, Kähönen M, Raitakari OT, Karakas M, Koenig W, Boehm BO, Winkelmann BR, März W, Lehtimäki T. Genome-wide association study pinpoints a new functional apolipoprotein B variant influencing oxidized low-density lipoprotein levels but not cardiovascular events: AtheroRemo Consortium. *Circ Cardiovasc Genet.* 2013 Feb;6(1):73-81. doi: 10.1161/CIRCGENETICS.112.964965. Epub 2012 Dec 17. PubMed PMID: 23247145.

6. Global Lipids Genetics Consortium, Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Do R, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkilä K, Hyppönen E, Isaacs A, Jackson AU, Johansson A, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikäinen LP, Magnusson PK, Mangino M, Mihailov E, Montasser ME, Müller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney AS, Döring A, Elliott P, Epstein SE, Eyjolfsson GI, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJ, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimäki T, Lin SY, Lindström J, Loos RJ, Mach F, McArdle WL, Meisinger C, Mitchell BD, Müller G, Nagaraja R, Narisu N, Nieminen TV, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruukonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stancáková A, Stirrups K, Swift AJ, Tietz L, Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YD, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrières J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllenstein U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Järvelin MR, Jula A, Kähönen M, Kaprio J, Kesäniemi A, Kivimäki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, März W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njølstad I, Pedersen NL, Power C,

- Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PE, Sheu WH, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BH, Ordovas JM, Boerwinkle E, Palmer CN, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Kathiresan S, Mohlke KL, Ingelsson E, Abecasis GR. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013 Nov;45(11):1274-83. doi: 10.1038/ng.2797. Epub 2013 Oct 6. PubMed PMID: 24097068; PubMed Central PMCID: PMC3838666.
7. Lange LA, Hu Y, Zhang H, Xue C, Schmidt EM, Tang ZZ, Bizon C, Lange EM, Smith JD, Turner EH, Jun G, Kang HM, Peloso G, Auer P, Li KP, Flannick J, Zhang J, Fuchsberger C, Gaulton K, Lindgren C, Locke A, Manning A, Sim X, Rivas MA, Holmen OL, Gottesman O, Lu Y, Ruderfer D, Stahl EA, Duan Q, Li Y, Durda P, Jiao S, Isaacs A, Hofman A, Bis JC, Correa A, Griswold ME, Jakobsdottir J, Smith AV, Schreiner PJ, Feitosa MF, Zhang Q, Huffman JE, Crosby J, Wassel CL, Do R, Franceschini N, Martin LW, Robinson JG, Assimes TL, Crosslin DR, Rosenthal EA, Tsai M, Rieder MJ, Farlow DN, Folsom AR, Lumley T, Fox ER, Carlson CS, Peters U, Jackson RD, van Duijn CM, Uitterlinden AG, Levy D, Rotter JI, Taylor HA, Gudnason V Jr, Siscovick DS, Fornage M, Borecki IB, Hayward C, Rudan I, Chen YE, Bottinger EP, Loos RJ, Sæthrom P, Hveem K, Boehnke M, Groop L, McCarthy M, Meitinger T, Ballantyne CM, Gabriel SB, O'Donnell CJ, Post WS, North KE, Reiner AP, Boerwinkle E, Psaty BM, Altshuler D, Kathiresan S, Lin DY, Jarvik GP, Cupples LA, Kooperberg C, Wilson JG, Nickerson DA, Abecasis GR, Rich SS, Tracy RP, Willer CJ; NHLBI Grand Opportunity Exome Sequencing Project. Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol. *Am J Hum Genet.* 2014 Feb 6;94(2):233-45. doi: 10.1016/j.ajhg.2014.01.010. PubMed PMID:24507775; PubMed Central PMCID: PMC3928660.
8. Brautbar A, Leary E, Rasmussen K, Wilson DP, Steiner RD, Virani S. Genetics of familial hypercholesterolemia. *Curr Atheroscler Rep.* 2015 Apr;17(4):491. doi: 10.1007/s11883-015-0491-z. Review. PubMed PMID: 25712136.
9. Awan Z, Choi HY, Stitzel N, Ruel I, Bamimore MA, Husa R, Gagnon MH, Wang RH, Peloso GM, Hegele RA, Seidah NG, Kathiresan S, Genest J. APOE p.Leu167del mutation in familial hypercholesterolemia. *Atherosclerosis.* 2013 Dec;231(2):218-22. doi: 10.1016/j.atherosclerosis.2013.09.007. Epub 2013 Sep 19. PubMed PMID: 24267230.
10. Futema M, Plagnol V, Li K, Whittall RA, Neil HA, Seed M; Simon Broome Consortium, Bertolini S, Calandra S, Descamps OS, Graham CA, Hegele RA, Karpe F, Durst R, Leitersdorf E, Lench N, Nair DR, Soran H, Van Bockxmeer FM; UK10K Consortium, Humphries SE. Whole exome sequencing of familial hypercholesterolaemia patients negative for LDLR/APOB/PCSK9 mutations. *J Med Genet.* 2014 Aug;51(8):537-44. doi: 10.1136/jmedgenet-2014-102405. Epub 2014 Jul 1. PubMed PMID: 24987033; PubMed Central PMCID: PMC4112429.
11. Radhakrishnan A, Ikeda Y, Kwon HJ, Brown MS, Goldstein JL. Sterol-regulated transport of SREBPs from endoplasmic reticulum to Golgi: oxysterols block transport by binding to Insig. *Proc Natl Acad Sci U S A.* 2007 Apr 17;104(16):6511-8. Epub 2007 Apr 11. PubMed PMID: 17428920; PubMed Central PMCID: PMC1851665.
12. Do R, Stitzel NO, Won HH, Jørgensen AB, Duga S, Angelica Merlini P, Kiezun A, Farrall M, Goel A, Zuk O, Guella I, Asselta R, Lange LA, Peloso GM, Auer PL; NHLBI Exome Sequencing Project, Girelli D, Martinelli N, Farlow DN, DePristo MA, Roberts R, Stewart AF, Saleheen D, Danesh J, Epstein SE, Sivapalaratnam S, Hovingh GK, Kastelein JJ, Samani NJ, Schunkert H, Erdmann J, Shah SH, Kraus WE, Davies R, Nikpay M, Johansen CT, Wang J, Hegele

RA, Hechter E, Marz W, Kleber ME, Huang J, Johnson AD, Li M, Burke GL, Gross M, Liu Y, Assimes TL, Heiss G, Lange EM, Folsom AR, Taylor HA, Olivieri O, Hamsten A, Clarke R, Reilly DF, Yin W, Rivas MA, Donnelly P, Rossouw JE, Psaty BM, Herrington DM, Wilson JG, Rich SS, Bamshad MJ, Tracy RP, Cupples LA, Rader DJ, Reilly MP, Spertus JA, Cresci S, Hartiala J, Tang WH, Hazen SL, Allayee H, Reiner AP, Carlson CS, Kooperberg C, Jackson RD, Boerwinkle E, Lander ES, Schwartz SM, Siscovick DS, McPherson R, Tybjaerg-Hansen A, Abecasis GR, Watkins H, Nickerson DA, Ardisino D, Sunyaev SR, O'Donnell CJ, Altshuler D, Gabriel S, Kathiresan S. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature*. 2015 Feb 5;518(7537):102-6. doi: 10.1038/nature13917. Epub 2014 Dec 10. PubMed PMID: 25487149; PubMed Central PMCID: PMC4319990.

13. Girelli D, Martinelli N, Peyvandi F, Olivieri O. Genetic architecture of coronary artery disease in the genome-wide era: implications for the emerging "golden dozen" loci. *Semin Thromb Hemost*. 2009 Oct;35(7):671-82. doi: 10.1055/s-0029-1242721. Epub . Review. PubMed PMID: 20013534.

14. Futema M, Plagnol V, Whittall RA, Neil HA; Simon Broome Register Group, Humphries SE; UK10K. Use of targeted exome sequencing as a diagnostic tool for Familial Hypercholesterolaemia. *J Med Genet*. 2012 Oct;49(10):644-9. doi: 10.1136/jmedgenet-2012-101189. PubMed PMID: 23054246; PubMed Central PMCID: PMC3475071.

15. Wild PS, Zeller T, Schillert A, Szymczak S, Sinning CR, Deiseroth A, Schnabel RB, Lubos E, Keller T, Eleftheriadis MS, Bickel C, Rupprecht HJ, Wilde S, Rossmann H, Diemert P, Cupples LA, Perret C, Erdmann J, Stark K, Kleber ME, Epstein SE, Voight BF, Kuulasmaa K, Li M, Schäfer AS, Klopp N, Braund PS, Sager HB, Demissie S, Proust C, König IR, Wichmann HE, Reinhard W, Hoffmann MM, Virtamo J, Burnett MS, Siscovick D, Wiklund PG, Qu L, El Mokthari NE, Thompson JR, Peters A, Smith AV, Yon E, Baumert J, Hengstenberg C, März W, Amouyel P, Devaney J, Schwartz SM, Saarela O, Mehta NN, Rubin D, Silander K, Hall AS, Ferrieres J, Harris TB, Melander O, Kee F, Hakonarson H, Schrezenmeir J, Gudnason V, Elosua R, Arveiler D, Evans A, Rader DJ, Illig T, Schreiber S, Bis JC, Altshuler D, Kavousi M, Witteman JC, Uitterlinden AG, Hofman A, Folsom AR, Barbalic M, Boerwinkle E, Kathiresan S, Reilly MP, O'Donnell CJ, Samani NJ, Schunkert H, Cambien F, Lackner KJ, Tiret L, Salomaa V, Munzel T, Ziegler A, Blankenberg S. A genome-wide association study identifies LIPA as a susceptibility gene for coronary artery disease. *Circ Cardiovasc Genet*. 2011 Aug 1;4(4):403-12. doi: 10.1161/CIRCGENETICS.110.958728. Epub 2011 May 23. PubMed PMID: 21606135; PubMed Central PMCID: PMC3157552.

16. Stitzel NO, Fouchier SW, Sjouke B, Peloso GM, Moscoso AM, Auer PL, Goel A, Gigante B, Barnes TA, Melander O, Orho-Melander M, Duga S, Sivapalaratnam S, Nikpay M, Martinelli N, Girelli D, Jackson RD, Kooperberg C, Lange LA, Ardisino D, McPherson R, Farrall M, Watkins H, Reilly MP, Rader DJ, de Faire U, Schunkert H, Erdmann J, Samani NJ, Charnas L, Altshuler D, Gabriel S, Kastelein JJ, Defesche JC, Nederveen AJ, Kathiresan S, Hovingh GK; National Heart, Lung, and Blood Institute GO Exome Sequencing Project. Exome sequencing and directed clinical phenotyping diagnose cholesterol ester storage disease presenting as autosomal recessive hypercholesterolemia. *Arterioscler Thromb Vasc Biol*. 2013 Dec;33(12):2909-14. doi: 10.1161/ATVBAHA.113.302426. Epub 2013 Sep 26. PubMed PMID: 24072694; PubMed Central PMCID: PMC4002172.

17. Damani SB, Topol EJ. Emerging genomic applications in coronary artery disease. *JACC Cardiovasc Interv*. 2011 May;4(5):473-82. doi: 10.1016/j.jcin.2010.12.016. Review. PubMed PMID: 21596318.

18. Musunuru K, Pirruccello JP, Do R, Peloso GM, Guiducci C, Sougnez C, Garimella KV, Fisher S, Abreu J, Barry AJ, Fennell T, Banks E, Ambrogio L, Cibulskis K, Kernysky A, Gonzalez E, Rudzicz N, Engert JC, DePristo MA, Daly MJ, Cohen JC, Hobbs HH, Altshuler D, Schonfeld G, Gabriel SB, Yue P,

- Kathiresan S. Exome sequencing, ANGPTL3 mutations, and familial combined hypolipidemia. *N Engl J Med*. 2010 Dec 2;363(23):2220-7. doi: 10.1056/NEJMoa1002926. Epub 2010 Oct 13. PubMed PMID: 20942659; PubMed Central PMCID: PMC3008575.
19. Zelcer N, Hong C, Boyadjian R, Tontonoz P. LXR regulates cholesterol uptake through Idol-dependent ubiquitination of the LDL receptor. *Science*. 2009 Jul 3;325(5936):100-4. doi: 10.1126/science.1168974. Epub 2009 Jun 11. PubMed PMID: 19520913; PubMed Central PMCID: PMC2777523.
20. Howles PN, Hui DY; Editorial Board. Physiological role of hepatic NPC1L1 in human cholesterol and lipoprotein metabolism: new perspectives and open questions. *J Lipid Res*. 2012 Nov;53(11):2253-5. doi: 10.1194/jlr.E031823. Epub 2012 Aug 31. PubMed PMID: 22941774; PubMed Central PMCID: PMC3465994.
21. Christoffersen M, Tybjaerg-Hansen A. Novel genes in LDL metabolism – a comprehensive overview. *Curr Opin Lipidol*. 2015 Jun;26(3):179-87. doi: 10.1097/MOL.0000000000000175. PubMed PMID: 25887678.
22. The Myocardial Infarction Genetics Consortium Investigators. Inactivating Mutations in *NPC1L1* and Protection from Coronary Heart Disease. *N Engl J Med*. Author manuscript; available in PMC 2015 May 27. Published in final edited form as: *N Engl J Med*. 2014 Nov 27; 371(22): 2072–2082. Published online 2014 Nov 12. doi: 10.1056/NEJMoa1405386

Tabella 3. Mutazioni LOF e *missenso* (7 out 7 software) identificate dopo l'analisi dei geni. La tabella mostra il numero di casi e di controlli portatori delle varianti in studio, la relativa sostituzione amminoacidica, l'effetto fenotipico atteso e il tipo di variazione. Solo relativamente alle mutazioni *missenso* è stato indicato anche se la variante era precedentemente associata (x) o meno (-) con i livelli di colesterolo.

Gene	Casi (n)	Controlli (n)	Posizione*	Sostituzione amminoacidica	Direzione**	Tipo***	Letteratura****
ABCG5	1	0	chr2:44041643_C/A	Glu579Ter	↑	LOF	
	1	0	chr2:44047211_G/A	Arg498Ter	↑	LOF	
	1	0	chr2:44055146_C/T	Ala204Thr	↑	missenso	-
	1	0	chr2:44055181_C/A	Gly192Val	↑	missenso	-
	0	1	chr2:44055202_A/T	Ile185Thr	↑	missenso	-
	1	0	chr2:44064975_G/C	Ser88Ter	↑	LOF	
	3	2	chr2:44065755_G/A	Gln22Ter	↑	LOF	
ABCG8	0	1	chr2:44071744_C/A	Tyr54Ter	↑	LOF	
	0	1	chr2:44078851_G/A	Val151Met	↑	missenso	-
	1	0	chr2:44078890_C/T	Arg164Ter	↑	LOF	
	0	1	chr2:44079575_G/T	Gly215Val	↑	missenso	-
	3	2	chr2:44099233_G/A	Trp361Ter	↑	LOF	

	1	0	chr2:44100948_C/T	Arg412Ter	↑	LOF	
	2	0	chr2:44101577_AC/A	Leu482TrpfsTer40	↑	LOF	
APOB	1	0	chr2:21225090_C/A	Glu4402Ter	↓	LOF	
	1	0	chr2:21225765_G/A	Arg4177Ter	↓	LOF	
	0	1	chr2:21227163_C/T	Trp4022Ter	↓	LOF	
	0	1	chr2:21229161_G/A	Arg3527Trp	↑	missenso	X
APOE	1	1	chr19:45412379_C/T	Gln276Ter	↑	LOF	
	0	2	chr19:45412386_G/GCC	Leu279ProfsTer	↑	LOF	
CYP7A1	1	0	chr8:59404257_TTTAAC/T	Leu430ValfsTer35	↑	LOF	
	0	1	chr8:59404983_G/A	Arg382Ter	↑	LOF	
CYP7A1	1	0	chr8:59409428_CTT/C	Lys214SerfsTer51	↑	LOF	
	0	2	chr8:59412572_ACTTAC/A	_*****	↑	LOF	
INSIG2	1	1	chr2:118864724_G/T	Gly199Ter	↑	LOF	
	0	1	chr2:118864751_C/T	Arg208Ter	↑	LOF	
LDLR	1	0	chr19:11210928_C/T	Gln33Ter	↑	LOF	
	0	1	chr19:11210974_G/A	Gly48Asp	↑	missenso	-
	1	0	chr19:11213359_CG/C	Asp72ThrfsTer134	↑	LOF	
	1	0	chr19:11213390_C/T	Arg81Cys	↑	missenso	X

	1	0	chr19:11213453_C/T	Gln102Ter	↑	LOF	
	5	0	chr19:11213463_G/A	_*****	↑	LOF	
	1	0	chr19:11215974_A/G	Asp131Gly	↑	missenso	-
	0	1	chr19:11215991_G/A	Gly137Ser	↑	missenso	X
	1	0	chr19:11215992_G/T	Gly137Val	↑	missenso	-
	2	0	chr19:11216000_G/T	Glu140Ter	↑	LOF	
	1	0	chr19:11216047_C/A	Cys155Ter	↑	LOF	
	1	0	chr19:11216112_C/T	Ser177Leu	↑	missenso	X
	1	0	chr19:11216171_T/C	Cys197Arg	↑	missenso	X
	8	0	chr19:11216244_A/G	Asp221Gly	↑	missenso	X
	1	0	chr19:11216247_G/A	Cys222Tyr	↑	missenso	X
	1	0	chr19:11218077_G/C	Cys276Ser	↑	missenso	-
	2	1	chr19:11221334_A/G	Asn316Ser	↑	missenso	X
	0	1	chr19:11221414_G/A	Gly343Ser	↑	missenso	X
	1	0	chr19:11221435_C/T	Arg350Ter	↑	LOF	
	1	1	chr19:11223962_G/A	Ala399Thr	↑	missenso	X
	1	0	chr19:11224013_C/T	Arg416Trp	↑	missenso	X
	1	0	chr19:11224024_C/G	Tyr419Ter	↑	LOF	

LDLR	2	0	chr19:11224061_C/G	Leu432Val	↑	missenso	X
	0	1	chr19:11224066_C/G	Asp433Glu	↑	missenso	-
	1	0	chr19:11224126_G/A	_*****	↑	LOF	
	1	0	chr19:11224326_G/A	Asp492Asn	↑	missenso	X
	1	0	chr19:11224419_G/A	Val523Met	↑	missenso	X
	1	0	chr19:11224422_G/A	Val524Met	↑	missenso	-
	0	1	chr19:11224428_C/T	Pro526Ser	↑	missenso	X
	2	0	chr19:11224437_G/C	Gly529Arg	↑	missenso	-
	1	0	chr19:11224439_G/A	_*****	↑	LOF	
	1	0	chr19:11226781_G/A	Trp533Ter	↑	LOF	
	1	0	chr19:11227549_C/T	Arg574Cys	↑	missenso	X
	2	1	chr19:11227604_G/A	Gly592Glu	↑	missenso	X
	2	0	chr19:11227613_G/A	Arg595Gln	↑	missenso	X
	1	0	chr19:11227676_T/C	_*****	↑	LOF	
	1	0	chr19:11230858_C/A	Leu646Ile	↑	missenso	-
	0	1	chr19:11230888_C/A	His656Asn	↑	missenso	X
	2	0	chr19:11231112_C/T	Pro685Leu	↑	missenso	X
	1	1	chr19:11231159_G/A	Gly701Ser	↑	missenso	X

	1	0	chr19:11234017_C/T	Gln770Ter	↑	LOF	
	2	0	chr19:11240210_T/TG	Val806GlyfsTer11	↑	LOF	
LDLRAP1	2	0	chr1:25883729_C/CA	His144GlnfsTer27	↑	LOF	
LIPA	1	0	chr10:91005469_G/A	Arg65Ter	↑	LOF	
ABO	0	2	chr9:136137573_T/C	_*****	↓	LOF	
ANGPTL3	0	1	chr1:63063580_AAT/A	Met116ValfsTer3	↓	LOF	
	0	2	chr1:63063592_GAACTC/G	Asn121LysfsTer3	↓	LOF	
	3	2	chr1:63063667_CAACT/C	Asn147Ter	↓	LOF	
ANGPTL3	2	0	chr1:63068052_G/GT	_*****	↓	LOF	
	1	0	chr1:63069787_T/A	Leu360Pro	↓	missenso	-
MYLIP	0	1	chr6:16143314_G/T	Leu176Phe	↓	missenso	-
	0	1	chr6:16146982_A/C	Ter446TyrextTer22	↓	LOF	
NPC1L1	2	0	chr7:44555477_G/A	Arg1268Cys	↓	missenso	-
	0	1	chr7:44556361_C/T	Gly1181Ser	↓	missenso	-
	0	2	chr7:44556462_A/G	Leu1147Pro	↓	missenso	-
	1	0	chr7:44561410_A/C	Ser952Ala	↓	missenso	-
	0	1	chr7:44573047_C/G	Asp798His	↓	missenso	-

	2	2	chr7:44573143_T/C	Met766Val	↓	missenso	-
	1	1	chr7:44575543_G/A	Arg627Cys	↓	missenso	-
	0	1	chr7:44576452_C/G	Gly557Ala	↓	missenso	-
	1	0	chr7:44578575_G/A	Pro474Leu	↓	missenso	-
	1	2	chr7:44578780_G/A	Arg406Ter	↓	LOF	
	1	2	chr7:44578861_G/T	Pro379Thr	↓	missenso	-
	0	3	chr7:44579108_AG/A	Ala296ValfsTer58	↓	LOF	
	0	1	chr7:44579378_C/A	Gln206His	↓	missenso	-
	1	0	chr7:44579506_A/G	Phe164Leu	↓	missenso	-
	1	0	chr7:44579575_G/A	Arg141Cys	↓	missenso	-
	1	0	chr7:44579613_G/A	Thr128Met	↓	missenso	-
	0	2	chr7:44579826_C/T	Cys57Tyr	↓	missenso	-
PCSK9	35	51	chr1:55505647_G/T	Arg46Leu	↓	missenso	X
	2	1	chr1:55512275_G/A	Arg160Gln	↓	missenso	-
	0	1	chr1:55518049_G/A	Val208Met	↓	missenso	-
	2	1	chr1:55523187_G/A	Gly394Ser	↓	missenso	-

	1	3	chr1:55523831_G/T	_*****	↓	LOF	
TM6SF2	1	0	chr19:19381141_G/A	Arg104Ter	↓	LOF	

* Posizione nucleotidica riferita al genoma umano di riferimento GRCh37/hg19, Febbraio 2009;

** ↑ la variante aumenta i livelli di colesterolo, ↓ la variante riduce i livelli di colesterolo;

*** LOF = mutazione *loss-of-function*: mutazione nonsense, inserzioni o delezioni *out-of-frame*, e mutazioni di *splicing*;

**** recuperate da: LOVD (Leiden Open-source Variation Database) [Fokkema et al., 2011], ExPASy (Expert Protein Analysis System) [Gasteiger et al., 2003], and ClinVar [Landrum et al., 2015];

***** queste varianti alterano i siti donatori o accettori di splicing.

Tabella 4 . Soggetti portatori di più di una mutazione nei geni sottoposti ad analisi

Caso/controllo	Varianti	Gene	Stato
Caso	chr19:11210928_C/T	<i>LDLR</i>	Omozigote
Caso	chr7:44578780_G/A chr7:44578575_G/A	<i>NPC1L1</i>	Eterozigote Eterozigote
Controllo	chr19:45412386_G/GCC chr1:63063592_GAACTC/G	<i>APOE</i> <i>ANGPTL3</i>	Eterozigote Eterozigote
Caso	chr7:44575543_G/A chr19:11226781_G/A	<i>NPC1L1</i> <i>LDLR</i>	Eterozigote Eterozigote
Caso	chr2:44078890_C/T chr19:11216244_A/G	<i>ABCG8</i> <i>LDLR</i>	Eterozigote Eterozigote
Controllo	chr2:44078851_G/A chr2:44065755_G/A	<i>ABCG8</i> <i>ABCG5</i>	Eterozigote Eterozigote

Tabella 5. Dataset finale delle mutazioni identificate nei casi e nei controlli

Geni	Varianti <i>loss-of-function</i> (LOF)		Varianti missenso (7out7)	
	Casi (n)	controlli (n)	Casi (n)	controlli (n)
ABCG5	6	2	2	1
ABCG8	7	3	0	2
APOB	2	2	0	0
APOE	1	3	0	0
CYP7A1	2	3	0	0
INSIG2	1	2	0	0
LDLR	48	8	7	2
LDLRAP1	2	0	0	0
LIPA	1	0	0	0
ABO	0	2	0	0
ANGPTL3	5	5	1	0
MYLIP	0	1	0	1
NPC1L1	2	5	10	13
PCSK9	36	54	4	3
TM6SF2	1	0	0	0

Tabella. 6 Caratteristiche cliniche e angiografiche stratificate sulla base della presenza delle varianti genetiche predisponenti l'ipercolesterolemia o delle varianti protettive l'ipocolesterolemia.

	PAZIENTI PORTATORI DELLE VARIANTI GENETICHE PREDISPONENTI (n=76)	PAZIENTI PORTATORI DELLE VARIANTI GENETICHE PROTETTIVE (n=57)	PAZIENTI PRIVI DELLE VARIANTI GENETICHE SELEZIONATE (n=1875)
ETÀ, MEDIA (DS)	38,68 (5,24)	39,14 (4,96)	39,66 (4,78)
• SESSO MASCHILE • SESSO FEMMINILE	67/76 (88, 1%) 11/76 (14,5%)	54/57 (94,7%) 3/75 (4%)	1666/1875 (88,8%) 209/1875 (11,2%)
FAMILIARITÀ PER CARDIOPATIA ISCHEMICA	67/76 (88,1%)	44/57 (77,2%)	1526/1875 (81,4%)
IPERTENSIONE ARTERIOSA	15/76 (19,7%)	15/57 (26,3%)	522/1875 (27,8%)
DIABETE MELLITO	7/76 (9,2%)	4/57 (7,0%)	143/1875 (7,6%)
COLESTEROLO, MEDIA (DS)	282 (86)	183 (44)	220 (52,6)
FUMATORI EX-FUMATORI	23/76 (30,3%) 43/76 (56,6%)	29/57 (50,9%) 20/57 (35,1%)	832/1875 (44,4%) 805/1875 (42,9%)
ESERCIZIO FISICO • OCCASIONALE • ABITUALE	15/76 (19,7%) 20/76 (26,3%)	11/57 (19,3%) 15/57 (26,3%)	422/1375 (30,8%) 481/1875 (25,6%)
ALCOOL	48/76 (63,1%)	31/57 (54,4%)	1147/1875 (61,2%)
BMI, MEDIA (DS)	26,03 (4,04)	27,62 (4,59)	26,7 (4,12)
CORONAROPATIA • MONOVASALE • MULTIVASALE	33/76 (43,4%) 36/76 (47,4%)	35/57 (61,4%) 9/57 (15,8%)	850/1875 (45,3%) 648/1875 (34,6%)
DUKE CORONARY SCORE, MEDIANA (RI)	37 (±24)	32 (±9)	32 (±19)

BIBLIOGRAFIA

1. Mathers CD, Salomon JA, Ezzati M, Begg S, Hoorn SV, Lopez AD. Global Burden of Disease and Risk Factors. New York, NY: Oxford University Press, 2006.
2. RSPP - Relazione sullo Stato Sanitario del Paese (RSSP) 2009-2010; https://www.salute.gov.it/imgs/C_17_pubblicazioni_1655_allegato.pdf
3. Palmieri L, Bennett K, Giampaoli S, Capewell S. Explaining the Decrease in Coronary Heart Disease Mortality in Italy between 1980 and 2000. *Am J Public Health* 2010; 99: 1-9
4. Sun T, Dong YH, Du W, Shi CY, Wang K, Tariq MA, Wang JX, Li PF. The Role of MicroRNAs in Myocardial Infarction: From Molecular Mechanism to Clinical Application. *Int J Mol Sci*. 2017 Mar 31;18(4). pii: E745. doi: 10.3390/ijms18040745.
5. Gimbrone MA Jr, García-Cardena G. Vascular endothelium, hemodynamics, and the pathobiology of atherosclerosis. *Cardiovasc Pathol*. 2016;22:9–15. doi: 10.1016/j.carpath.2012.06.006.
6. Stary HC. Natural history and histological classification of atherosclerotic lesions: an update. *Arterioscler Thromb Vasc Biol*. 2000;20:1177–1178.
7. Virmani R, Kolodgie FD, Burke AP, Farb A, Schwartz SM. Lessons from sudden coronary death: a comprehensive morphological classification scheme for atherosclerotic lesions. *Arterioscler Thromb Vasc Biol*. 2000;20:1262–1275.
8. Simionescu N, Vasile E, Lupu F, Popescu G, Simionescu M. Prelesional events in atherogenesis. Accumulation of extracellular cholesterol-rich liposomes in the arterial intima and cardiac valves of the hyperlipidemic rabbit. *Am J Pathol*. 1986;123:109–125.
9. Libby P. Mechanisms of acute coronary syndromes. *N Engl J Med*. 2013;369:883–884. doi: 10.1056/NEJMc1307806.
10. Quillard T, Araújo HA, Franck G, Shvartz E, Sukhova G, Libby P. TLR2 and neutrophils potentiate endothelial stress, apoptosis and detachment: implications for superficial erosion. *Eur Heart J*. 2015;36:1394–1404. doi:10.1093/eurheartj/ehv044
11. Grundy SM, Pasternak R, Greenland P, Smith S, Jr, Fuster V. Assessment of cardiovascular risk by use of multiple-risk-factor assessment equations: a statement for healthcare professionals from the American Heart Association and the American College of Cardiology. *Circulation*. 1999;100:1481–1492. doi: 10.1161/01.CIR.100.13.1481
12. Smith SC Jr, Blair SN, Criqui MH, Fletcher GF, Fuster V, Gersh BJ, Gotto AM, Gould L, Greenland P, Grundy SM, Hill MN, Hlatky MA, Houston-Miller N, Krauss RM, LaRosa J, Ockene IS, Oparil S, Pearson TA, Rapaport E, Starke R, and Secondary Prevention Panel.

- Preventing heart attack and death in patients with coronary disease. *Circulation*. 1995;92:2–4.
13. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837–1847
 14. Berg JN, Tymoczko JL, Stryer L. *BIOCHIMICA*, sesta edizione, 2008.
 15. Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, Hegele RA, Krauss RM, Raal FJ, Schunkert H, Watts GF, Borén J, Fazio S, Horton JD, Masana L, Nicholls SJ, Nordestgaard BG, van de Sluis B, Taskinen MR, Tokgözoğlu L, Landmesser U, Laufs U, Wiklund O, Stock JK, Chapman MJ, Catapano AL. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur Heart J*. 2017 Aug 21;38(32):2459–2472. doi: 10.1093/eurheartj/ehx144.
 16. Brautbar A, Leary E, Rasmussen K, Wilson DP, Steiner RD, Virani S. Genetics of familial hypercholesterolemia. *Curr Atheroscler Rep*. 2015 Apr;17(4):491. doi: 10.1007/s11883-015-0491-z. Review
 17. Nordestgaard B, Chapman M, Humphries S, et al. Familial hypercholesterolemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease. *Eur Heart J*. 2013;34(45):3478–90. This is the Consensus Statement of the European atherosclerosis Society discussing the underdiagnoses of FH, its genetics, and screening approach.
 18. Seidah NG, Awan Z, Chrétien M, Mbikay M. PCSK9: a key modulator of cardiovascular health. *Circ Res*. 2014;114:1022–36. A thorough review of the biological and genetic characteristics of PCSK9 including animal models and the possible clinical utility
 19. Cuchel M, Bruckert E, Ginsberg H, et al. Homozygous familial hypercholesterolaemia: new insights and guidance for clinicians to improve detection and clinical management. A position paper from the Consensus Panel on Familial Hypercholesterolaemia of the European Hypercholesterolaemia of the European Atherosclerosis Society. *Eur Heart J*. 2014;35:2146–57. This is the Consensus Statement of the European atherosclerosis Society and provides an updated approach to genetics and screening of familial hypercholesterolemia
 20. Tada H, Kawashiri M, Ohtani R, et al. A novel type of familial hypercholesterolemia: double heterozygous mutations in LDL receptor and LDL receptor adaptor protein 1 gene. *Atherosclerosis*. 2011;219:663–6.
 21. Marduel M, Carrie A, Sassolas A, et al. Molecular spectrum of autosomal dominant hypercholesterolemia in France. *Hum Mutat*. 2010;31:E1811–1824.
 22. Ejarque I, Real J, Martinez-Hervas S, et al. Evaluation of clinical diagnosis criteria of familial ligand defective apoB 100 and lipoprotein phenotype comparison between LDL receptor gene mutations affecting ligand-binding domain and the R3500Q mutation of the apoB gene in patients from a South European population. *Transl Res*. 2008;151:162–7.

23. Ahmad Z, Adams-Huet B, Chen C, Garg A. Low prevalence of mutations in known loci for autosomal dominant hypercholesterolemia in multi-ethnic patient cohort. *Circ Cardiovasc Genet.* 2012;5:666–75.
24. Alonso R, Mata N, Castillo S, et al. Cardiovascular disease in familial hypercholesterolaemia: influence of low-density lipoprotein receptor mutation type and classic risk factors. *Atherosclerosis.* 2008;200:315–21.
25. Hopkins P, Toth P, Ballantyne C, Rader D. Familial hypercholesterolemias: prevalence, genetics, diagnosis and screening recommendations from the National Lipid Association Expert Panel on Familial Hypercholesterolemia. *J Clin Lipidol.* 2011;5:S9–17.
26. <http://www.hgmd.cf.ac.uk/ac/gene.php?gene=LDLR>
27. http://www.ncbi.nlm.nih.gov/clinvar?LinkName=gene_clinvar&from_uid=3949
28. http://www.ucl.ac.uk/ldlr/Current/index.php?select_db=LDLR
29. Hobbs H, Russell D, Brown M, et al. The LDL receptor locus and familial hypercholesterolemia: mutational analysis of a membrane protein. *Annu Rev Genet.* 1990;24:133–70.
30. Hobbs H, Brown M, Goldstein J. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat.* 1992;1:445–66.
31. Whitfield A, Barrett H, van Bockxmeer F, Burnett J. Lipid disorder and mutations in the APOB gene. *Clin Chem.* 2004;50:1725–32.
32. Blackhart B, Ludwig E, Pierotti V, et al. Structure of the human apolipoprotein B gene. *J Biol Chem.* 1986;261:15364–7.
33. Myant N. Familial defective apolipoprotein B-100: a review, including some comparisons with familial hypercholesterolaemia. *Atherosclerosis.* 1993;104:1–18.
34. Boren J, Ekstrom U, Agren B, Nilsson-Ehle P, Innerarity T. The molecular mechanism for the genetic disorder familial defective apolipoprotein B100. *J Biol Chem.* 2001;276:9214–8.
35. Rader D, Cohen J, Hobbs H. Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. *J Clin Invest.* 2003;111(12):1795–803.
36. Innerarity T, Mahley R, Weisgraber K, et al. Familial defective apolipoprotein B-100: a mutation of apolipoprotein B that causes hypercholesterolemia. *J Lipid Res.* 1990;31:1337–49.
37. Soufi M, Sattler A, Maerz W, et al. A new but frequent mutation of APOB-100 APOB His 3543Tyr. *Atherosclerosis.* 2004;174:11–6.

38. Motazacker MM, Pirruccello J, Huijgen R, et al. Advances in genetics show the need for extending screening strategies for autosomal dominant hypercholesterolaemia. *Eur Heart J*. 2012;33:1360–6.
39. Thomas ER, Atanur SS, Norsworthy PJ, et al. Identification and biochemical analysis of a novel APOB mutation that causes autosomal dominant hypercholesterolemia. *Mol Genet Genomic Med*. 2013;1:155–61.
40. Cohen J, Boerwinkle E, Mosley Jr T, Hobbs H. Sequence variations in PCSK9, low LDL and protection against coronary heart disease. *N Engl J Med*. 2006a;354:1264–72.
41. Cohen , J. C. , E. Boerwinkle , T. H. Mosley , Jr ., and H. H. Hobbs . 2006 . Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N. Engl. J. Med*. **354** : 1264 – 1272 .
42. Guella I, Asselta R, Ardissino D, Merlini PA, Peyvandi F, Kathiresan S, Mannucci PM, Tubaro M, Duga S. Effects of PCSK9 genetic variants on plasma LDL cholesterol levels and risk of premature myocardial infarction in the Italian population. *J Lipid Res*. 2010;51:3342–3349
43. Turgeon RD, Pearson GJ. Proprotein convertase subtilisin/kexin type 9 inhibitors for reduction of cardiovascular events. *Am J Health Syst Pharm*. 2018 Jun 1;75(11):747-754. doi: 10.2146/ajhp170707.
44. Pisciotta L, Oliva C, Pes G, et al. Autosomal recessive hypercholesterolemia (ARH) and homozygous familial hypercholesterolemia (FH): a phenotypic comparison. *Atherosclerosis*. 2006a;188:398–405.
45. Mabuchi H, Nohara A, Noguchi T, et al. Genotypic and phenotypic features in homozygous familial hypercholesterolemia caused by proprotein convertase subtilisin/kexin type 9 (PCSK9) gain-of-function mutation. *Atherosclerosis*. 2014;236:54–61.
46. Soutar A, Naoumova R. Mechanisms of disease: genetic causes of familial hypercholesterolemia. *Nat Clin Pract Cardiovasc Med*. 2007;4:214–25.
47. Rahalkar A, Hegele R. Monogenic pediatric dyslipidemias: classification, genetics and clinical spectrum. *Mol Genet Metab*. 2008;93:282–94.
48. Filigheddu F, Quagliarini F, Campagna F, et al. Prevalence and clinical features of heterozygous carriers of autosomal recessive hypercholesterolemia in Sardinia. *Atherosclerosis*. 2009;207:162–7.
49. Garcia CK, Wilund K, Arca M, et al. Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. *Science* 2001;292: 1394-8.
50. Naoumova R, Neuwirth C, Lee P, et al. Autosomal recessive hypercholesterolaemia: long-term follow up and response to treatment. *Atherosclerosis*. 2004;174:165–72.

51. Myers RH, Kiely DK, Cupples LA, Kannel WB. Parental history is an independent risk factor for coronary artery disease: the Framingham Study. *Am Heart J*. 1990 Oct;120(4):963-9.
52. Schildkraut JM, Myers RH, Cupples LA, Kiely DK, Kannel WB. Coronary risk associated with age and sex of parental heart disease in the Framingham Study. *Am J Cardiol*. 1989 Sep 15;64(10):555-9.
53. Shea S, Ottman R, Gabrieli C, Stein Z, Nichols A. Family history as an independent risk factor for coronary artery disease. *J Am Coll Cardiol*. 1984 Oct;4(4):793-801.
54. European Cardiovascular Disease statistics. 2009 edition.
55. Roberts R. Genetics of coronary artery disease. *Circ. Res*. 2014;114:1890–1903. doi: 10.1161/CIRCRESAHA.114.302692.
56. Roberts R, Wells GA, Stewart AF, Dandona S, Chen L. The genomewide association study—a new era for common polygenic disorders. *J Cardiovasc Transl Res*. 2010;3:173–182
57. Jarcho JA, McKenna W, Pare JA, Solomon SD, Holcombe RF, Dickie S, Levi T, Donis-Keller H, Seidman JG, Seidman CE. Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. *N Engl J Med*. 1989;321:1372–1378.
58. Hejtmancik JF, Brink PA, Towbin J, Hill R, Brink L, Tapscott T, Trakhtenbroit A, Roberts R. Localization of gene for familial hypertrophic cardiomyopathy to chromosome 14q1 in a diverse US population. *Circulation*. 1991;83:1592–1597
59. Geisterfer-Lowrance A, Kass S, Tanigawa G, et al. A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosium heavy chain missense mutation. *Cell*. 1990;62:999–1006
60. Wang Q, Shen J, Splawski I, Atkinson D, Li Z, Robinson JL, Moss AJ, Towbin JA, Keating MT. SCN5A mutations associated with an inherited cardiac arrhythmia, long QT syndrome. *Cell*. 1995;80:805–811.
61. Chen Q, Kirsch GE, Zhang D, et al. Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. *Nature*. 1998;392:293–296
62. Brugada R, Tapscott T, Czernuszewicz GZ, Marian AJ, Iglesias A, Mont L, Brugada J, Girona J, Domingo A, Bachinski LL, Roberts R. Identification of a genetic locus for familial atrial fibrillation. *N Engl J Med*. 1997;336:905–911.
63. Olson TM, Michels VV, Thibodeau SN, Tai YS, Keating MT. Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science*. 1998;280:750–752.
64. Gollob MH, Green MS, Tang AS, Gollob T, Karibe A, Ali Hassan AS, Ahmad F, Lozado R, Shah G, Fananapazir L, Bachinski LL, Roberts R, Hassan AS. Identification of a gene responsible for familial Wolff- Parkinson-White syndrome. *N Engl J Med*. 2001;344:1823–1831

65. Li D, Ahmad F, Gardner MJ, Weilbaecher D, Hill R, Karibe A, Gonzalez O, Tapscott T, Sharratt GP, Bachinski LL, Roberts R. The locus of a novel gene responsible for arrhythmogenic right-ventricular dysplasia characterized by early onset and high penetrance maps to chromosome 10p12-p14. *Am J Hum Genet.* 2000;66:148–156.
66. Kruglyak L, Nickerson DA. Variation is the spice of life. *Nat Genet.* 2001;27:234–236.
67. Kruglyak L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet.* 1999;22:139–144.
68. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet.* 2005;6:95–108
69. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet.* 2005;6:109–118.
70. Morgan TM, Krumholz HM, Lifton RP, Spertus JA. Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. *JAMA.* 2007;297:1551–1561.
71. Anand S, Xie C, Pare G, Montpetit A, Rangarajan S, McQueen M, Cordell HJ, Keavney B, Yusef S, Hudson TJ, Engert JC. Genetic variants associated with myocardial infarction risk factors in over 8,000 individuals from five ethnic groups: INTERHEART Study. *Circ Cardiovas Genet.* 2009;2:16–25
72. LaFramboise T. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Res.* 2009;37:4181–4193
73. Gibbs RA. The International HapMap Project. *Nature.* 2003;426:789–796
74. Stranger BE, Forrest MS, Dunning M, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science.* 2007;315:848–853.
75. Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. *Cell.* 2012 Mar 16;148(6):1242-57. doi: 10.1016/j.cell.2012.03.001. Review.
76. Musunuru K, Pirruccello JP, Do R, Peloso GM, Guiducci C, Sougnez C, Garimella KV, Fisher S, Abreu J, Barry AJ, Fennell T, Banks E, Ambrogio L, Cibulskis K, Kernysky A, Gonzalez E, Rudzicz N, Engert JC, DePristo MA, Daly MJ, Cohen JC, Hobbs HH, Altshuler D, Schonfeld G, Gabriel SB, Yue P, Kathiresan S. Exome sequencing, ANGPTL3 mutations, and familial combined hypolipidemia. *N Engl J Med.* 2010 Dec 2;363(23):2220-7. doi: 10.1056/NEJMoa1002926. Epub 2010 Oct 13. PubMed PMID: 20942659; PubMed Central PMCID: PMC3008575.
77. Christoffersen, Mette; Tybjærg-Hansen, Anne. Novel genes in LDL metabolism – a comprehensive overview. *Curr Opin Lipidol.* 2015 Jun;26(3):179-87. doi: 10.1097/MOL.0000000000000175.

78. The Myocardial Infarction Genetics Consortium Investigators. Inactivating Mutations in NPC1L1 and Protection from Coronary Heart Disease. *N Engl J Med*. Author manuscript; available in PMC 2015 May 27. Published in final edited form as: *N Engl J Med*. 2014 Nov 27; 371(22): 2072–2082. Published online 2014 Nov 12. doi: 10.1056/NEJMoa1405386
79. Sorrentino V, Fouchier SW, Motazacker MM, Nelson JK, Defesche JC, Dallinga-Thie GM, Kastelein JJ, Kees Hovingh G, Zelcer N. Identification of a loss-of-function inducible degrader of the low-density lipoprotein receptor variant in individuals with low circulating low-density lipoprotein. *Eur Heart J*. 2013 May;34(17):1292-7. doi: 10.1093/eurheartj/ehs472. Epub 2013 Jan 16.
80. Chan L, Boerwinkle E. Gene-environment interactions and gene therapy in atherosclerosis. *Cardiol Rev*. 1994;2:130–137
81. Deloukas P, Kanoni S, Willenborg C, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*. 2013;45:25–33.
82. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461:747–753
83. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *J. Am. Med. Assoc* 2008;299:1335–1344
84. Maher B. Personal genomes: The case of the missing heritability. *Nature* 2008;456:18–21. [PubMed: 18987709]
85. Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001;17:502–510.[PubMed: 11525833]
86. Collins FS, Guyer MS, Chakravarti A. Variations on a theme: cataloging human DNA sequence variation. *Science* 1997;278:1580–1581. [PubMed: 9411782]
87. Pritchard JK. Are rare variants responsible for susceptibility to common diseases? *Am. J. Hum. Genet*. 2001;69:124–137. [PubMed: 11404818]
88. McCarthy MI, Hirschhorn JN. Genome-wide association studies: potential next steps on a genetic journey. *Hum. Mol. Genet* 2008a;17(R2):R156–R165. [PubMed: 18852205] Insightful review of initial findings from GWAS, the heritability that they do and do not explain, and potential for progress from other GWAS, identification of rare variants, and studies of epigenetics and gene expression and function.
89. McCarthy MI, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Rev. Genet* 2008b;9:356–369. [PubMed: 18398418]
90. Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends Genet* 2008;24:133–141. [PubMed: 18262675]

91. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* 2015 Jan;43(Database issue):D789-98. doi: 10.1093/nar/gku1205. Epub 2014 Nov 26. PubMed PMID: 25428349; PubMed Central PMCID: PMC4383985.
92. Liu X, Jian X, Boerwinkle E. 2011. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum Mutat* 32:894–899.
93. Kumar P, Henikoff S, Ng PC. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4:1073–1081.
94. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–249
95. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. 2010. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* 7:575–576.
96. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* 2011; 39(17): e118
97. Chun S, Fay JC. 2009. Identification of deleterious mutations within three human genomes. *Genome Res* 19:1553–1561
98. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GL, Edwards KJ, Day IN, Gaunt TR. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Hum Mutat.* 2013;34(1):57-65.
99. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat.* 2011 May;32(5):557-63. doi: 10.1002/humu.21438. Epub 2011 Feb 22. PubMed PMID: 21520333.
100. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003 Jul 1;31(13):3784-8. PubMed PMID: 12824418; PubMed Central PMCID: PMC168970
101. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, Jang W, Katz K, Ovetsky M, Riley G, Sethi A, Tully R, Villamarin-Salomon R, Rubinstein W, Maglott DR. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 2015 Nov 17. pii: gkv1222. [Epub ahead of print] PubMed PMID: 26582918
102. Varret M, Rabès JP, Collod-Bérout G, Junien C, Boileau C, Bérout C. Software and database for the analysis of mutations in the human LDL receptor gene. *Nucleic Acids Res.* 1997 Jan 1;25(1):172-80. PubMed PMID: 9016531; PubMed Central PMCID: PMC146377.

103. Mark DB, Nelson CL, Califf RM, Harrell FE Jr, Lee KL, Jones RH, Fortin DF, Stack RS, Glower DD, Smith LR, et al. Continuing evolution of therapy for coronary artery disease. Initial results from the era of coronary angioplasty. *Circulation*. 1994 May;89(5):2015-25.
104. Tg, et al. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N Engl J Med*. 2014;371:22–31. doi: 10.1056/NEJMoa1307095
105. Guella I, Asselta R, Ardissino D, Merlini PA, Peyvandi F, Kathiresan S, Mannucci PM, Tubaro M, Duga S. Effects of PCSK9 genetic variants on plasma LDL cholesterol levels and risk of premature myocardial infarction in the Italian population. *J Lipid Res*. 2010;51:3342–3349
106. Members of the Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia. Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *N Engl J Med*. 2016;374(19):1898 doi: 10.1056/NEJMxx160012
107. Do R, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature*. 2015;518:102–106. doi: 10.1038/nature13917.
108. Mega JL, Stitzel NO, Smith JG, Chasman DI, Caulfield MJ, Devlin JJ, Nordio F, Hyde CL, Cannon CP, Sacks FM, Poulter NR, Sever PS, Ridker PM, Braunwald E, Melander O, Kathiresan S, Sabatine MS. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet*. 2015 Jun 6;385(9984):2264-71. doi: 10.1016/S0140-6736(14)61730-X