- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.223.124.244 Mon, 29 Apr 2024, 13:36:54

Protocollo di studio

Il PROCARDIS: un approccio attuale allo studio della genetica dell'infarto miocardico

Simona Barlera, Benedetta Diamante Chiodini, Maria Grazia Franzosi, Gianni Tognoni

Dipartimento di Ricerca Cardiovascolare, Istituto di Ricerche Farmacologiche Mario Negri, Milano

Key words:
Epidemiology; Genetics; Methodology; Myocardial infarction; Statistics.

Coronary artery disease is a complex and multifactorial pathology. Although the environmental component of coronary artery disease has been thoroughly investigated and is hence well known, knowledge about the genetic factors implicated in this disease is still scarce.

Technological advances and the fact that the Human Genome Project has almost been completed allow the application of approaches that were not feasible few years ago to the genetic investigation of complex diseases.

The aim of the PROCARDIS study is to identify new susceptibility genes of precocious coronary artery disease, through a genome-wide screen applying statistical methods of linkage analysis followed by a family-based association study.

The originality of PROCARDIS lies in the fact that it is an international multicenter study. This allows recruitment of a very large number of individuals so that the population size, considered up to now unachievable, is adequate for the aims of the study.

(Ital Heart J Suppl 2001; 2 (9): 997-1004)

© 2001 CEPI Srl

Ricevuto il 19 giugno 2001; accettato il 4 luglio 2001.

Per la corrispondenza:

Dr.ssa Simona Barlera

Dipartimento di Ricerca Cardiovascolare
Istituto di Ricerche Farmacologiche
Mario Negri
Via Eritrea, 62
20157 Milano
E-mail: simo@marionegri.it

Introduzione

Dall'analisi della letteratura sui principali geni candidati alla malattia coronarica, quali ad esempio quelli coinvolti nel metabolismo lipidico, nel sistema emostatico, nel metabolismo dell'omocisteina e nel sistema renina-angiotensina, sono emerse evidenze per lo più negative o poco conclusive[1]. Si afferma perciò sempre più la necessità di impiegare un approccio metodologico diverso all'indagine genetica della cardiopatia ischemica che superi i limiti degli studi di associazione e sia in grado di valutare anche geni finora non "candidati", ossia non sospettati di un coinvolgimento nei processi patogenetici della malattia.

A partire da queste premesse, il presente scritto comprende:
- un quadro di riferimento delle tecniche statistiche e degli approcci dell'epidemiologia genetica che sono specificamente necessari per affrontare le indagini non centrate sui geni candidati;
- la presentazione della proposta PROCARDIS (PRecOcious Coronary ARtery DISease), dove dovrebbe essere possibile riconoscere il tentativo fatto da un consorzio di ricerca collaborativa internazionale di rispettare le regole del gioco di una ricerca capace di produrre risultati affidabili;
- un glossario dei termini tecnici che si incontrano sempre più nella letteratura e nella pratica cardiologica

Quadro di riferimento

Gli avanzamenti tecnologici in biologia molecolare, statistica e informatica occorsi negli ultimi due decenni (Tab. I), si sono rivelati determinanti per l'identificazione dei geni di suscettibilità a malattie ereditarie, consentendo oggi di applicare alla genetica umana metodologie di analisi classica, il cui utilizzo è stato per lungo tempo limitato proprio per la mancanza di strumenti e tecnologie adeguati.

Il vero punto di svolta è consistito nell'opportunità di riscontrare direttamente le variazioni del DNA umano (polimorfismi) e di utilizzarle quali marcatori genetici per la mappatura di loci responsabili di malattie ereditarie attraverso l'analisi di linkage.

L'analisi di linkage. L'analisi di linkage rappresenta una metodologia statistica classica dell'analisi genetica, che mira a stabilire una correlazione tra la trasmissione del gene-malattia in famiglie di pa-

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.223.124.244 Mon, 29 Apr 2024, 13:36:54

Tabella I. Principali tappe dell'evoluzione in genetica negli ultimi 50 anni.

1950-1955	Ricostruzione dei componenti biochimici e della struttura chimica a doppia elica del DNA; scoperta della DNA polimerasi, enzima capace di replicare il DNA, e degli enzimi batterici (di restrizione) che tagliano il DNA.
1975-1979	Inizio del sequenziamento del DNA; isolamento dei primi geni nell'uomo, e messa in luce dell'esistenza di considerevoli variazioni (polimorfismi) nelle sequenze nucleotidiche del genoma umano.
1980-1985	Inizio della cartografia genomica tramite l'utilizzo di polimorfismi di lunghezza dei frammenti di restrizione; sviluppo di tecniche di amplificazione genica, in particolare della reazione polimerasica a catena.
1987 →	Identificazione dei minisatelliti VNTRS; prima mappa pubblicata dell'intero genoma.
1989 →	Identificazione dei microsatelliti VNTRS; prende ufficialmente il via lo Human Genome Project.
2000 →	Il consorzio Human Genome Project e l'azienda Celera Genomics annunciano il completamento della sequenza grezza del genoma umano.
Febbraio 2001	Nature (15 Feb 2001, 409: 6822) e Science (16 Feb 2001, 291: 5507) pubblicano gli studi sulla sequenza pressoché integrale del genoma umano

VNTRS = variable nucleotide tandem repeat sequences.

zienti affetti e la trasmissione di marcatori polimorfici di cui sia nota la localizzazione sul genoma[2]. La sua applicazione prevede generalmente l'indagine di marcatori disposti sull'intero genoma umano (genome-wide screen). La base biologica è il concetto di ricombinazione o crossing-over, grazie al quale durante la prima fase meiotica avviene lo scambio fisico di materiale genetico tra i cromosomi omologhi paterno e materno (Fig. 1).

Se tra due loci avviene un crossing-over, allora gli alleli che il figlio riceve a tali loci dal genitore, si dicono "ricombinati" non essendo più identici a quelli trasmessi dal genitore stesso. È naturale che più i loci sono vicini tra loro sul cromosoma, minore è la probabilità che tra essi avvenga la ricombinazione, e di conseguenza maggiore sarà la probabilità che gli alleli a tali loci vengano trasmessi insieme[3]. Quest'ultimo fenomeno, in contraddizione alla legge di Mendel sulla segregazione indipendente dei loci, è noto come "linkage genetico". In caso di linkage si osserva una frequenza di ricombinazione (θ), intesa come percentuale dei gameti che ci si attende siano ricombinanti (< 50%), e tanto più piccola quanto più i loci sono vicini. L'obiettivo dell'analisi di linkage consiste per l'appunto nello stimare la frequenza di ricombinazione (misura della distanza genetica tra due loci), e quindi nello stabilire se questa è significativamente < 50% (Tab. II).

Sebbene i principi del linkage fossero già noti alla fine degli anni '50, la sua applicazione si è resa possibile, e in questi anni si sta rivelando via via più agevole, grazie alla continua scoperta di marcatori sempre più numerosi, polimorfici e informativi (polimorfismi di lunghezza dei frammenti di restrizione, microsatelliti, polimorfismi di singoli nucleotidi, ecc.).

I momenti successivi di un'analisi sull'intero genoma. Nella ricerca dei geni di suscettibilità alla malattia, l'analisi di linkage rappresenta il primo passo per l'identificazione delle regioni subcromosomiche candidate a contenerli. La seconda fase consiste nella saturazione delle regioni individuate attraverso l'utilizzo di ulteriori marcatori allo scopo di restringere maggiormente l'area candidata. Poiché la risoluzione della mappatura genetica ottenuta attraverso l'analisi di linkage è limitata, non potendo scendere al di sotto di 1 cM (pari ad una distanza fisica di 1000 chilobasi), la seconda fase prevede generalmente l'analisi del "linkage disequilibrium" tra il gene-malattia e gli alleli dei marcatori ad esso più strettamente concatenati. In questo stadio si va a valutare l'esistenza di un'associazione preferenziale tra determinati alleli dei marcatori polimorfici e il gene responsabile della malattia, attraverso uno studio di associazione. Esistono due tipi di studi di associazione: quello classico di popolazione,

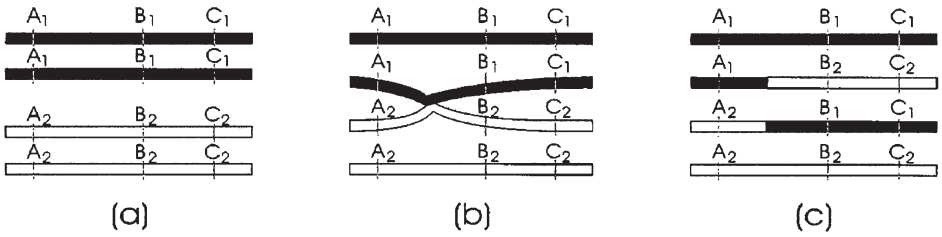

Figura 1. *Ricombinazione tra cromosomi omologhi durante la I fase meiotica. In* (a) *i due cromosomi si accoppiano e ciascun cromosoma (materno e paterno) si divide in due cromatidi identici. Sono indicati gli alleli appartenenti ai tre loci (A, B e C). In* (b) *si osserva un crossing-over tra i loci A e B. In* (c) *sono raffigurati i quattro gameti risultanti; in due è avvenuta una ricombinazione tra i loci A e B e tra i loci A e C; nessuna ricombinazione è avvenuta tra B e C.*

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.223.124.244 Mon, 29 Apr 2024, 13:36:54

Tabella II. L'analisi di linkage.

Con il termine "analisi di linkage" ci si riferisce ai metodi statistici per la localizzazione dei geni che contribuiscono alle patologie umane, utili nella fase di individuazione della regione subcromosomica iniziale (candidata). Il fenomeno del linkage (concatenazione) si deve alla stretta vicinanza di due geni o marcatori lungo un cromosoma e si verifica qualora la cosegregazione tra un marcatore (oggi per lo più microsatelliti) ed un gene che influenza il fenotipo di interesse si manifesta con una frequenza maggiore a quella attesa per il puro effetto del caso.

Analisi di linkage tradizionale (parametrica)
I metodi di linkage tradizionali (generalmente applicati ai tratti mendeliani) necessitano la preventiva specificazione del modello di ereditarietà che spiega la trasmissione della malattia nelle famiglie. Vengono messe a confronto due ipotesi: H_0: la frequenza di ricombinazione (θ) tra il gene di suscettibilità alla malattia ed un dato marcatore è pari al 50%, cioè i due loci non sono linked; H_1: la frequenza di ricombinazione (θ) è < 50%, cioè una specifica regione cromosomica contiene un gene di suscettibilità alla malattia, che si trova accanto (linked) ad un marcatore genetico noto. La forza dell'evidenza in favore del linkage viene valutata attraverso il rapporto delle funzioni di verosimiglianza (LR) e convenzionalmente misurata tramite la statistica del LOD score. Quest'ultimo, è il logaritmo su base 10 del rapporto di verosimiglianza tra l'ipotesi di linkage (H_1) e l'ipotesi di indipendenza (H_0).
Perciò: LR = Probabilità (Dati$|0 \leq \theta < 0.5$)/Probabilità (Dati$|\theta = 0.5$)

$$\text{LOD score} = \log_{10} (\text{LR})$$

La stima di massima verosimiglianza è quella che genera il più alto LOD score (maximum LOD score). L'ipotesi di linkage è considerata statisticamente più verosimile dell'ipotesi di non linkage quando il LOD score supera una soglia critica, generalmente LOD ≥ 3. Questo valore indica che i dati osservati sono 1000 volte più probabili assumendo l'ipotesi di linkage rispetto a quella di non linkage. Se i due loci non si trovano vicini tra loro e dunque sono non-linked, allora i LOD score saranno per lo più negativi; per convenzione LOD score più bassi di -2 sono considerati quale evidenza di non linkage.

Analisi di linkage non parametrica
I metodi di linkage non parametrici sono considerati la migliore strategia per localizzare geni (di solito più di uno) coinvolti nella suscettibilità alle malattie complesse, essendo indipendenti dalla specificazione del modello di trasmissione. Un comune approccio consiste nella stima della condivisione allelica tra coppie di parenti affetti allo scopo di determinare se soggetti malati per una certa patologia e imparentati ereditano copie di alleli "identici per discendenza" (IBD) più spesso di quanto ci si aspetterebbe per puro effetto del caso. Le coppie di fratelli affetti costituiscono la struttura familiare più semplice alla quale si applicano i metodi di condivisione allelica. Questo approccio si basa sulla premessa che se due fratelli sono entrambi affetti da una certa malattia, essi hanno con tutta probabilità ereditato dai loro genitori copie identiche di una data regione cromosomica (contenente l'allele di suscettibilità alla malattia). Due fratelli condividono in media per ciascun locus zero, uno o due alleli IBD con probabilità del 25, 50 e 25% rispettivamente. In un campione di coppie di fratelli affetti, in assenza di linkage tra un marcatore e il gene malattia, ci si attendenderà dunque una simile distribuzione, mentre in caso di linkage si osserverà un aumento delle coppie di fratelli che condividono uno o due alleli IBD (ed una diminuzione delle coppie che condividono zero alleli IBD). L'evidenza di linkage può essere stimata: a) comparando il numero osservato ed atteso di coppie che condividono zero, uno o due alleli identici per stato (IBS) mediante un test χ^2 con due gradi di libertà; b) tramite il calcolo statistico della massima verosimiglianza (MLS). Quest'ultimo valuta l'incremento di condivisione allelica tra parenti affetti, e il presupposto per accettare l'ipotesi di linkage è che la statistica MLS sia > 3.

nel quale la frequenza allelica (di un determinato locus) viene confrontata tra casi e controlli tra loro non imparentati, e quello su base familiare, che include il "transmission disequilibrium test" (TDT), che è il più frequentemente utilizzato (Tab. III). Il gene responsabile della malattia viene infine isolato e caratterizzato funzionalmente solo nella fase ultima del clonaggio posizionale[4] (Tab. IV).

Lo studio delle malattie complesse. Negli ultimi 10 anni, l'utilizzo dell'analisi di linkage ha portato a successi straordinari nell'identificazione dei geni responsabili di parecchie malattie monogeniche (mendeliane), generando inoltre la speranza di poter affrontare anche tratti più comuni e complessi.

La tradizionale analisi di linkage, cosiddetta "parametrica", si rivela tuttavia più adatta allo studio di malattie semplici a trasmissione mendeliana, richiedendo

Tabella III. Transmission disequilibrium test.

Il "transmission disequilibrium test" costituisce il test di associazione su base familiare più frequentemente utilizzato, nel quale i soli campioni necessari sono quelli dell'individuo affetto (proband) e dei suoi genitori (trios). Il gruppo dei controlli è costituito dagli alleli non trasmessi dai genitori ai figli affetti e il campione dei casi dagli alleli trasmessi e quindi presenti nei proband. Mediante questo test di associazione si esamina quante volte i genitori eterozigoti per un certo allele trasmettono l'allele in questione al figlio affetto e quante volte non lo trasmettono. Se la trasmissione dell'allele avviene con frequenza significativamente > 50%, ciò costituisce evidenza di associazione allelica e linkage tra marcatore e malattia. La significatività statistica viene calcolata attraverso l'utilizzo della statistica del χ^2 (McNamar's test). L'applicazione di uno studio di associazione su base familiare è consigliabile in quanto evita il problema della disomogeneità tra casi e controlli, caratteristico degli studi di associazione di popolazione.

Tabella IV. Clonaggio posizionale.

Clonaggio posizionale significa isolare il gene partendo unicamente dalla localizzazione subcromosomica, senza utilizzare alcuna informazione riguardante la patogenesi o la funzione biochimica. Questa regione iniziale, identificata mediante l'utilizzo di tecniche di analisi dell'intero genoma e di linkage in famiglie affette dalla patologia di interesse, definisce generalmente un'area relativamente ampia (10 Mb o più). La strategia del clonaggio posizionale consiste nel costruire le mappe fisica e genetica della regione, precisare la localizzazione subcromosomica, e quindi identificare i geni presenti per valutarli come possibili geni patologici. Per restringere la regione candidata sono necessarie serie di cloni e di marcatori all'interno della regione, e oggi sono disponibili marcatori polimorfici distanti non più di 1 cM. Il clonaggio posizionale rimane nonostante tutto una strategia ardua, e probabilmente in futuro sempre meno necessaria. Gli avanzamenti del Progetto Genoma Umano e il conseguente aumento di geni mappati in specifiche regioni subcromosomiche permetteranno un approccio sempre più posizionale al gene candidato. Una volta mappata la patologia sarà cioè sempre più possibile effettuare una semplice ricerca nelle banche dati per identificare i geni candidati.

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.223.124.244 Mon, 29 Apr 2024, 13:36:54

la specificazione di un preciso modello genetico, che include parametri quali il modello di trasmissione, la frequenza dell'allele-malattia, e la penetranza di ciascun genotipo. Viceversa tale strategia risulta piuttosto inappropriata di fronte alla maggioranza delle malattie complesse, data la difficoltà di specificare il modello genetico sottostante. Appartengono a questo gruppo tutte le malattie comuni, quali l'ipertensione, il diabete, il cancro, diverse malattie neurologiche e psichiatriche, e certamente le malattie cardiovascolari, nelle quali si pensa siano coinvolti più geni in combinazione con diversi fattori ambientali.

Quindi, per evitare errori che possono derivare dall'impiego di modelli genetici non corrispondenti alla realtà è preferibile l'utilizzo di metodi statistici non parametrici, ossia che non richiedono la specificazione del modello di trasmissione della malattia. Il metodo non parametrico più comune, ovvero l'analisi su coppie di fratelli affetti (ASPs), si basa sul presupposto che se due fratelli sono entrambi ammalati della stessa patologia avranno molto probabilmente ereditato dai loro genitori copie identiche della stessa regione subcromosomica (contenente il gene di suscettibilità alla malattia)[5] (Tab. II). Dato l'interesse sempre maggiore che si sta sviluppando nei confronti dell'eziologia delle malattie multifattoriali, attualmente questo approccio è utilizzato frequentemente, ed è facile immaginare che le nuove opportunità dell'era post-genomica lo renderanno un metodo sempre più diffuso. Nella realtà la sua applicazione è apparsa fino ad oggi piuttosto ardua e deludente, a causa sia dell'enorme numero di coppie di fratelli necessario allo studio di questo genere di patologie, sia del relativamente scarso numero di marcatori disponibili. Un esempio valido in questo senso è lo studio pionieristico di Davies et al.[6] del 1994, nel quale il metodo ASPs fu applicato per la ricerca dei geni di suscettibilità al diabete di tipo 1 nell'ambito di un'indagine sull'intero genoma. Lo studio portò risultati importanti che confermarono la poligenicità della trasmissione della malattia, ma affinché fossero definitivi sarebbe stato necessario un campione decisamente più esteso.

Analoga è stata la più recente esperienza di Sharma et al.[7], che hanno effettuato una preliminare ricerca sistematica dei geni maggiori di suscettibilità all'ipertensione essenziale sull'intero genoma umano. Come per il diabete, l'indagine ha dimostrato come non esista una singola regione cromosomica responsabile dell'origine del disordine, e come la predisposizione all'ipertensione essenziale dipenda da più geni. I risultati dello studio dovranno comunque venire confermati da studi in grado di mappare più finemente le regioni candidate e di dimensioni sostanzialmente maggiori, affinché si possa passare all'ultima fase del clonaggio posizionale.

Più una patologia è multifattoriale, infatti, meno la componente genetica risulta influente sul piano eziologico, e di conseguenza il rischio relativo di sviluppare la malattia attribuibile ai soli fattori genetici si rivela modesto e inversamente proporzionale al numero di pazienti necessari per lo studio. Quindi malattie complesse, caratterizzate da un rischio relativo genetico difficilmente quantificabile ma certamente piuttosto contenuto, quali diabete, ipertensione e cardiopatia ischemica, richiedono una dimensione del campione considerata finora irrealistica sul piano pratico (per la malattia coronarica più di 2000 ASPs). Ad oggi non esistono risultati disponibili di studi di linkage su tutto il genoma miranti ad investigare le basi genetiche della cardiopatia ischemica.

Sintesi del protocollo dello studio PROCARDIS

Il PROCARDIS è il primo studio che si prefigge di attuare un'analisi di linkage su tutto il genoma, con l'obiettivo di individuare e caratterizzare nuovi geni che influenzano l'insorgenza della malattia coronarica precoce. Nella prima fase di "genome-wide screen", sarà applicato l'approccio di linkage non parametrico ASPs su almeno 2000 ASPs affetti da cardiopatia ischemica ad età < 65 anni, al fine di individuare eventuali regioni subcromosomiche candidate a contenere geni di suscettibilità alla malattia coronarica precoce. Una volta identificate le aree di potenziale interesse verrà applicato il TDT, al fine di restringerle. Le fasi successive del progetto prevedono la caratterizzazione fisica e funzionale sempre più precisa degli eventuali nuovi loci candidati attraverso tecniche di clonaggio posizionale, e infine la determinazione dell'impatto di questi geni sulla suscettibilità alla malattia coronarica nella popolazione generale.

In considerazione dell'ampio numero di famiglie necessario per avere concrete possibilità di individuare un linkage, lo studio PROCARDIS nasce come progetto internazionale di collaborazione multicentrica europea.

Allo studio PROCARDIS collaborano, infatti, cinque gruppi di ricerca che hanno la responsabilità degli aspetti scientifici: il Dipartimento di Medicina Cardiovascolare e l'Unità di Ricerche Cliniche dell'Università di Oxford, l'Istituto per la Ricerca sull'Aterosclerosi dell'Università di Münster, il King Gustav V Research Institute del Karolinska Hospital di Stoccolma e il GISSI (Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico). Le analisi biochimiche vengono effettuate principalmente in laboratori specializzati presso le Università di Oxford, Münster e Stoccolma. Oxagen (Abingdon, UK), una nuova compagnia biotecnologica, istituita grazie al supporto dell'Università di Oxford e del Wellcome Trust, si occupa della genotipizzazione.

Il braccio italiano dello studio si è dotato di un Comitato Etico e di un Comitato Scientifico che operano in modo autonomo in accordo alla linea internazionale del progetto (vedi Appendice).

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.223.124.244 Mon, 29 Apr 2024, 13:36:54

Metodi. Lo studio PROCARDIS si propone di applicare il metodo di linkage non parametrico ASPs nell'ambito di un'analisi su tutto il genoma.

A questo scopo verrà valutato se qualcuno tra l'ampio numero di marcatori analizzati sull'intero genoma è ereditato dai fratelli affetti all'interno delle famiglie con una frequenza maggiore di quella che ci si aspetterebbe per il puro effetto del caso. Nella prima fase del progetto verranno utilizzati 400 marcatori, con un intervallo medio di 10 cM ed un'eterozigosità media del 78.9%.

Questi marcatori potrebbero rivelarsi "linked" ai geni di suscettibilità alla malattia nelle regioni cromosomiche in cui si trovano questi ultimi, consentendo in tal modo l'individuazione delle aree subcromosomiche "candidate". Per raggiungere una potenza statistica sufficiente a mettere in luce un eventuale linkage è necessaria un'ampia raccolta di ASPs.

Per la valutazione della dimensione del campione necessaria si fa riferimento alla statistica λ_s, che dà una stima del contributo della genetica alla suscettibilità alla malattia coronarica. Lo studio del Swedish Twin Registry[8] ha riportato un rischio relativo (λ_s) di morte da infarto del miocardio del 3.8 tra i fratelli gemelli maschi dizigoti i cui co-gemelli erano anch'essi deceduti per la stessa malattia ad un'età < 55 anni (se paragonati a quelli il cui co-gemello non era morto). Perciò lo studio PROCARDIS si basa su un λ_s pari a 3.8 per la malattia coronarica. La dimensione del campione è inversamente proporzionale al valore della statistica λ_s. Poiché la malattia coronarica è caratterizzata da un valore di λ_s piuttosto modesto, anche quando paragonata ad altri tratti complessi quali il diabete mellito insulino-dipendente ($\lambda_s = 15$), la schizofrenia ($\lambda_s = 8.6$) e il diabete mellito non insulino-dipendente ($\lambda_s = 3.5$), è necessario un campione decisamente ampio. È stato calcolato che sono necessarie circa 2000 ASPs per avere una potenza dell'85% nell'identificare il linkage con un gene che contribuisce per il 10% all'intera componente genetica della malattia coronarica precoce (Fig. 2).

La strategia del PROCARDIS prevede un'analisi preliminare di primi 1000 ASPs. Sulla base dei risultati di questa prima fase, verrà quindi effettuata un'analisi completa o parziale del genoma sulle successive 1000 ASPs. Per l'analisi dei dati verrà applicato, a seconda dell'informatività dei nuclei familiari, l'approccio a più punti di "identità per discendenza" e/o di "identità per stato" (IBD/IBS), e verrà considerata quale evidenza di linkage la condizione in cui un marcatore o due marcatori vicini raggiungono il livello di significatività standard ($p < 0.05$) una volta corretto per il numero di marcatori testati. Le informazioni relative ai fenotipi intermedi saranno utili per capire se i linkage riscontrati possono essere spiegati dalla variazione di particolari marcatori biochimici, e ciò costituirebbe la chiave per individuare in un certo locus probabili geni candidati o vie coinvolte nella patogenesi. Allo scopo di indagare i potenziali linkage che emergono da queste analisi iniziali sul genoma, si passerà successivamente al mappaggio preciso di alcune regioni mediante una combinazione di analisi di linkage e studi di associazione. Pertanto saranno eseguite ulteriori analisi ASPs, e sarà inoltre effettuata un'analisi di associazione intrafamiliare, il TDT, sulle stesse e su altre famiglie che risulteranno informative per questo genere di approccio. In questa fase, si realizzerà un mappaggio di saturazione per ciascuna regione di potenziale linkage, utilizzando 30-50 marcatori microsatelliti (ad intervalli di circa 1 cM) identificati dai database di pubblico dominio. Questo approccio graduale dovrebbe consentire la definizione di regioni "candidate" di dimensioni compatibili con le tecniche di clonaggio posizionale, che sarà quindi applicato ai linkage più promettenti.

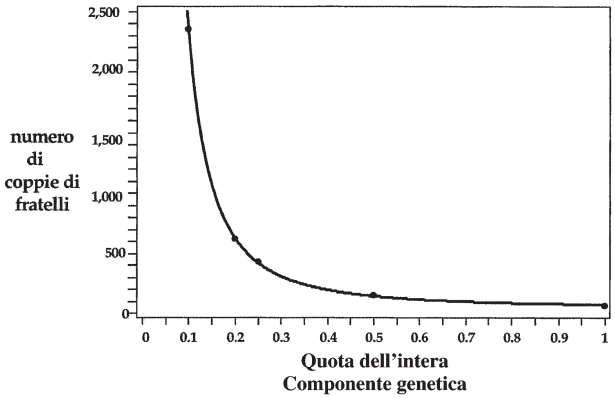

Figura 2. *Curva della dimensione del campione per individuare il linkage (potenza 85%, p = 0.001) per una malattia con rischio relativo (λ_s) pari a 3.8. Per mettere in luce loci che contribuiscono per il 10-15% alla componente genetica totale della malattia coronarica sono necessarie 2000-2500 coppie di fratelli affetti.*

I momenti chiave nell'isolamento del gene-malattia saranno: la costruzione delle mappe fisiche di specifiche regioni; la ricerca di polimorfismi nei geni candidati per la malattia; la conferma dell'identità dei geni-malattia mediante ulteriori analisi genetiche e studi funzionali.

Criteri di selezione. Sono necessarie diverse tipologie di famiglie in relazione agli approcci che si dovranno utilizzare.

Coppie di fratelli affetti. Famiglie costituite da due o più fratelli/sorelle che hanno sviluppato la malattia coronarica prima dei 65 anni, dando priorità a coloro che hanno sviluppato la malattia prima dei 55 anni, con o senza i genitori. Si cerca di reclutare tanti più fratelli/sorelle affetti e tanti più fratelli/sorelle/genitori non malati possibili, fino ad un massimo di 4 fratelli.

• Criteri di selezione per i componenti delle famiglie con coppie di fratelli affetti. Età: nello studio Swedish Twin, gli intervalli di confidenza per il rischio relativo stimato di evento fatale da infarto miocardico ad età differenti sono ampi, e dunque rimane incerta l'entità

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.223.124.244 Mon, 29 Apr 2024, 13:36:54

con la quale si indebolisce l'effetto genetico con il passare degli anni. Inoltre, con l'avanzare dell'età non solo aumenta in modo sostanziale la prevalenza di malattia coronarica, ma aumenta anche la prevalenza di "casi" con fratelli affetti. Ad esempio, nel database dell'International Study of Infarct Survival (ISIS), solo il 4% dei soggetti con infarto miocardico ad età < 55 anni risultava avere un fratello affetto, in confronto al 9% dei soggetti con infarto miocardico tra i 60 e i 64 anni[9]; risultati simili sono stati riscontrati anche dal GISSI[10]. Pertanto, è difficile predire quale sarà l'effetto sulla potenza statistica di un simultaneo incremento del limite superiore di età e del numero di ASPs in studi di linkage genetico. La strategia adottata dallo studio PROCARDIS è quella di reclutare un minimo di 2000 ASPs, a condizione che entrambi i fratelli abbiano sviluppato malattia coronarica prima dei 65 anni, e dando priorità alle coppie con prima diagnosi nel proband prima dei 55 anni.

• Criteri diagnostici per i proband: i soggetti con infarto miocardico documentato (secondo i criteri di diagnosi dell'Organizzazione Mondiale della Sanità) o altre sindromi coronariche acute sono eleggibili come proband. La diagnosi di infarto miocardico richiede la presenza di due o più criteri diagnostici tra:
- dolore toracico ischemico tipico, edema polmonare, sincope o shock;
- sviluppo di onda Q patologica e/o comparsa e scomparsa di sopraslivellamento del tratto ST seguita da inversione dell'onda T in due o più derivazioni elettrocardiografiche;
- incremento della concentrazione degli enzimi serici diagnostici di infarto del miocardio (ad esempio creatinfosfochinasi più del doppio rispetto ai valori normali).

Per confermare la diagnosi di sindrome coronarica acuta è necessaria la documentazione dell'avvenuta ospedalizzazione nel caso di una delle seguenti situazioni:
- angina instabile diagnosticata per il dolore toracico tipico ischemico e associata a depressione del tratto ST reversibile in due o più derivazioni elettrocardiografiche standard;
- trombolisi per sospetto infarto miocardico (indicato da sopraslivellamento del tratto ST in due o più derivazioni elettrocardiografiche standard) anche in assenza di successiva inversione dell'onda T, o comparsa di onda Q o significativo incremento enzimatico;
- rivascolarizzazione di emergenza (ad esempio, al momento del ricovero) in seguito a presentazione del tipico dolore toracico di natura ischemica.

• Criteri diagnostici per i fratelli affetti: i soggetti con infarto miocardico documentato o con altre sindromi coronariche acute sintomatiche o con diagnosi di angina cronica stabile sintomatica, fratelli di proband, sono eleggibili. L'angina cronica stabile è definita dalla presenza di dolore toracico tipico ischemico da sforzo, che si allevia con il riposo o con la nitroglicerina, confermata da un test da sforzo positivo o da altre procedure diagnostiche.

I soggetti sottoposti a rivascolarizzazione per il controllo dei sintomi sono eleggibili, mentre non lo sono coloro che non sono mai stati sintomatici.

Transmission disequilibrium test. Famiglie costituite da un soggetto maschio o femmina che ha sviluppato malattia coronarica prima dei 65 anni di età (proband) e da due genitori viventi, o da un genitore ed almeno un fratello vivente non affetto. Si cerca di reclutare quanti più fratelli è possibile fino ad un massimo di 4. È comunque indispensabile che venga reclutato almeno un genitore, e che venga reclutato un fratello/sorella nel caso sia disponibile solo un genitore.

• Selezione dei soggetti per le famiglie TDT. Per quanto riguarda età e criteri diagnostici dei proband valgono gli stessi principi illustrati per gli ASPs.

Aspetti etici. Il protocollo, i materiali informativi e il modulo di consenso del paziente a partecipare (formulato in accordo alle leggi vigenti in termini di trattamento dei dati personali) sono stati approvati dal Comitato Etico dello studio, composto da un cardiologo con competenze di bioetica, due clinici medici con competenze di metodologia della ricerca, un esperto in campo giuridico, un esperto in biometria, un esperto nell'ambito dell'informazione/tutela dei pazienti, un esperto nel settore della ricerca genetica. È previsto il consenso scritto del paziente e di ciascun familiare che accetta di collaborare al PROCARDIS.

Conclusioni

Lo studio PROCARDIS si prefigge di attuare un'analisi su tutto il genoma in un numero di famiglie sufficientemente ampio da consentire la scoperta di geni che influenzano l'insorgenza della malattia coronarica precoce. Il progetto si colloca in un momento ben preciso caratterizzato da un lato dalla necessità di chiarire e comprendere a fondo il ruolo della componente genetica nell'ambito di una malattia complessa quale la malattia coronarica, dall'altro dai progressi tecnologici e dall'ultimazione del Progetto Genoma Umano, che rendono perseguibili oggi approcci considerati inattuabili fino a pochi anni or sono.

I risultati di uno studio importante come il PROCARDIS potrebbero contribuire a creare nuove prospettive in campo preventivo e terapeutico[11]. Tuttavia, prima di immaginare sviluppi straordinari in campi oggi molto alla moda, quali la farmacogenetica e/o la terapia genica, è ragionevole limitarsi a traguardi che al momento attuale appaiono più plausibili e concreti. L'identificazione di suscettibilità genetiche individuali,

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.223.124.244 Mon, 29 Apr 2024, 13:36:54

probabilmente riconducibili non ad un unico gene, ma piuttosto a profili di rischio complessi prodotti da più varianti genetiche e dalla loro interazione con l'ambiente, potrebbe costituire un ulteriore indicatore di livello di rischio personale. Pertanto in quest'ottica la genetica contribuirebbe a potenziare l'efficienza delle carte del rischio nel costruire una prevenzione sempre più "aggiustata" sull'individuo.

Riassunto

Allo stato attuale delle conoscenze, non vi sono più dubbi circa la natura multifattoriale e complessa della malattia coronarica. Tuttavia, se la componente ambientale è stata studiata approfonditamente ed è attualmente piuttosto ben conosciuta, le nozioni acquisite riguardo alla componente genetica sono solo preliminari.

I progressi tecnologici e l'ultimazione del Progetto Genoma Umano rendono perseguibili oggi approcci di indagine genetica alle malattie complesse considerati inattuabili fino a pochi anni or sono, quali ad esempio l'analisi dell'intero genoma attraverso tecniche di linkage seguite da studi di associazione intrafamiliari.

Lo studio PROCARDIS si prefigge di attuare per la prima volta un'analisi su tutto il genoma in un numero di famiglie sufficientemente ampio da consentire la scoperta di geni che influenzano l'insorgenza della malattia coronarica precoce. La caratteristica di innovatività del PROCARDIS consiste nella sua natura multicentrica e internazionale, che consente il reclutamento di un numero di famiglie finora considerato di difficile realizzazione.

Parole chiave: Epidemiologia; Genetica; Infarto miocardico; Metodologia; Statistica.

Ringraziamenti

Gli autori ringraziano la Sig.ra Alessandra Carnaghi per il prezioso contributo nella redazione del manoscritto.

Appendice

Comitati e Segreterie dello studio PROCARDIS

International Project Steering Committee
- University of Oxford, Department of Cardiovascular Medicine, Wellcome Trust Centre for Human Genetics, UK: Fiona Green, Martin Farrall, Donna Harper, Elizabeth Taylor, Hugh Watkins (Chairman); Clinical Trial Service Unit: Rory Collins, Robert Clarke
- Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico (GISSI), Italia: Maria Grazia Franzosi, Gianni Tognoni
- Karolinska Hospital, Stoccolma, Svezia: Anders Hamsten, Mai-Lis Hellénius
- Universität Münster, Institut für Arterioskleroseforschung, Germania: Gerd Assmann, Helmut Schulte
- Oxagen Ltd, Abingdon, UK: Mark Edwards, Jenny Taylor
- Astra Zeneca, Mölndal, Svezia: Björn Löwenadler, Gunnar Olsson

Settore Italiano
- Comitato Scientifico GISSI-PROCARDIS: Eloisa Arbustini (ANMCO), Maria Grazia Franzosi (Presidente, IRFMN), Marco Tubaro (ANMCO), Gianni Tognoni (IRFMN)
- Comitato Etico: Albano Del Favero, Enrico Geraci (Presidente), Umberto Loi, Ettore Marubini, Gianna Milano, Sergio Ottolenghi, Luigi Pagliaro
- Segreteria Scientifica: Benedetta Chiodini, Luisella Crociati, Letizia Ferrario
- Statistica e Gestione Dati: Simona Barlera, Enrico Björn Nicolis
- Segreteria Organizzativa: Alessandra Carnaghi, Cristina Masini

Glossario

Alleli. Differenti forme di uno stesso gene che possono avere una diversa espressione fisica (sequenza del DNA).

Aplotipo. Una combinazione di alleli che si trovano in loci associati su un singolo cromosoma (paterno o materno).

ASPs (affected sibling pairs). Strategia di linkage non parametrica che utilizza coppie di fratelli/sorelle affetti/e. Confronta la percentuale di condivisione allelica esistente tra i fratelli ad un determinato locus rispetto alla percentuale attesa per il solo effetto del caso.

CentiMorgan (cM). Unità che esprime la distanza relativa tra i geni o marcatori collocati su uno stesso cromosoma, pari all'1% della frazione di ricombinazione; 1 cM corrisponde ad una distanza fisica di circa 1 megabase (pari a 1 milione di coppie di basi).

Dizigoti (gemelli). Due o più individui nati dallo stesso parto che provengono dalla fecondazione di due uova da parte di due spermatozoi differenti. Essi spartiscono soltanto il 50% del patrimonio genetico e sono pertanto assimilabili a fratelli.

Gene candidato (approccio). Procedura volta ad individuare il gene responsabile di un determinato difetto genetico. Il gene viene individuato tra quelli che mostrano caratteristiche strutturali o funzionali tali da renderli sospettabili di essere alla base del difetto genetico sotto indagine.

IBD (identity by descendent). Gli alleli sono definiti identici per discendenza qualora appaiano identici e si sa che essi provengono dallo stesso allele della generazione ancestrale di riferimento; risulta quindi tracciabile la trasmissione all'interno della famiglia. Nelle analisi di linkage gli alleli IBD si rivelano più informativi di quelli IBS.

IBS (identity by state). Gli alleli sono definiti identici per stato qualora appaiano identici, ma non è noto se essi provengano dallo stesso comune allele ancestrale. Gli studi che utilizzano coppie di fratelli affetti si devono spesso basare sull'identità IBS; quella IBD non può essere stabilita se i genitori non vengono reclutati nello studio e la trasmissione nelle famiglie non può essere tracciata.

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 18.223.124.244 Mon, 29 Apr 2024, 13:36:54

λ_s. Rischio di morte/malattia rispetto alla popolazione generale di un fratello qualora l'altro fratello sia deceduto/ammalato per la medesima patologia.

Linkage (concatenazione). Si parla di linkage in riferimento alla trasmissione di alleli su loci differenti. Tanto più due alleli si trovano distanti uno dall'altro, tanto maggiore è la probabilità che tra essi si verifichi un evento ricombinante. Tanto più due alleli si trovano vicini sullo stesso cromosoma, tanto minore è la probabilità che avvenga un evento ricombinante. In quest'ultima situazione gli alleli si comportano come geneticamente accoppiati e questo fenomeno è detto linkage genetico.

Linkage disequilibrium. Trasmissione non casuale genitori-prole di alleli di geni o di marcatori prossimi gli uni agli altri su di uno stesso cromosoma.

Locus (loci). Parte di DNA che segue le leggi di trasmissione mendeliana. Il gene è uno speciale tipo di locus, cioè è un locus con una specifica funzione (prodotto genico).

Marcatori genomici. Brevi sequenze nucleotidiche (DNA) note, più o meno variabili tra individui di una stessa popolazione, utili come marcatori del genoma. Numerosi tipi di sequenze sono considerati marcatori, tra i quali geni funzionali, porzioni di sequenze espresse, piccoli segmenti di DNA individuati tramite la reazione polimerasica a catena, microsatelliti, polimorfismi di lunghezza dei frammenti di restrizione e polimorfismi di singoli nucleotidi.

Mendeliana (eredità). Eredità di un carattere specificato da un singolo gene. Sinonimo di eredità monofattoriale.

Monozigoti (gemelli). Due o più individui nati dallo stesso parto che provengono dallo stesso uovo fecondato. Essi possiedono il medesimo patrimonio genetico.

Multipoint linkage methods. Metodi utili alla localizzazione dei geni che contribuiscono ai fenotipi complessi, che si rifanno (utilizzano) molteplici marcatori simultaneamente. Al contrario, i metodi di linkage a due punti (two-point) prendono in considerazione un marcatore alla volta per svelare il linkage con un gene che influenza il tratto di interesse.

Proband. Individuo affetto dalla patologia in oggetto che rappresenta il caso di riferimento nel reclutamento di un nucleo familiare.

Ricombinazione (crossing-over). Scambio reciproco di materiale genetico tra regioni corrispondenti nelle due copie (di origine materna e paterna) dello stesso cromosoma durante la fase meiotica.

Transmission disequilibriun test (TDT). Strategia che indaga la trasmissione dell'allele-malattia da un genitore eterozigote a un figlio affetto. Un particolare allele in linkage disequilibrium con l'allele-malattia verrà trasmesso più frequentemente al figlio affetto di quanto avverrebbe se tutti gli alleli venissero trasmessi in modo casuale.

Bibliografia

1. Chiodini BD, Barlera S, Franzosi MG, Tognoni G. I geni di suscettibilità all'infarto: una revisione della letteratura. Ital Heart J Suppl 2001; 2: 935-44.
2. Devoto M. Metodi statistici per la mappatura genetica di loci responsabili di malattie ereditarie a trasmissione mendeliana e complessa. In: Cao A, Durand P, eds. Genetica molecolare per il pediatra. Milano: CIS, 1997: 45-50.
3. Olson JM, Witte JS, Elston RC. Tutorial in biostatistics. Genetic mapping of complex traits. Stat Med 1999; 18: 2961-81.
4. Strachan T, Read AP. Genetica umana molecolare. Torino: UTET, 1996: 372.
5. Ellsworth DL, Manolio TA. The emerging importance of genetics in epidemiologic research III. Bioinformatics and statistical genetic methods. Ann Epidemiol 1999; 9: 207-24.
6. Davies JL, Kawaguchi Y, Bennet ST, et al. A genome-wide search for human type 1 diabetes susceptibility genes. Nature 1994; 371: 130-6.
7. Sharma P, Fatibene J, Ferraro F, et al. A genome-wide search for susceptibility loci to human essential hypertension. Hypertension 2000; 35: 1291-6.
8. Marenberg ME, Risch N, Berkman LF, et al. Genetic susceptibility to death from coronary heart disease in a study of twins. N Engl J Med 1994; 330: 1041-6.
9. Parish S, Collins R, Peto R, et al. Cigarette smoking, tar yields, and non-fatal myocardial infarction: 14 000 cases and 32 000 controls in the United Kingdom. The International Studies of Infarct Survival (ISIS) Collaborators. BMJ 1995; 311: 471-7.
10. Roncaglioni MC, Santoro L, D'Avanzo B, et al. Role of family history in patients with myocardial infarction. An Italian case-control study. Circulation 1992; 85: 2065-72.
11. Ellsworth D, Sholinsky P, Jaquish C, Fabsitz RR, Manolio TA. Coronary heart disease. At the interface of molecular genetics and preventive medicine. Am J Prev Med 1999; 16: 122-33.