

S.S. Epidemiologia Molecolare

Tecniche avanzate di analisi statistica e bioinformatica applicate a studi su biomarcatori in popolazioni umane

Linea di ricerca: 1 – Prevenzione e Cancerogenesi

Programma: b - Biomarcatori biologici e molecolari di esposizione, di danno, di suscettibilità e di rischio di cancro

Responsabile scientifico: Marcello Ceppi

Altro personale della struttura partecipante al progetto: Cecilia Lando, Monica Neri, Stefano Moretti, Fabio Gallo

Anno di inizio: 2009

Durata: 36 mesi

Parole chiave: analisi statistica; bioinformatica; epidemiologia molecolare; genomica; high throughput

Altre strutture IST partecipanti: S.C. Epidemiologia e Biostatistica (D. Ugolini); S.C. Epidemiologia Clinica (P. Bruzzi); S.S. Oncologia Traslazionale Pediatrica (G.P. Tonini)

Altri Enti coinvolti: University of Maastricht, Holland (J. Kleinjans); CSIRO, Adelaide, Australia (M. Fenech); Università di Pisa (R. Barale); Laboratorio di Biologia Molecolare, Istituto G. Gaslini, Genova (L. Varesio); Dipartimento di Matematica, Università di Genova (F. Patrone), IRCCS S. Raffaele Pisana, Roma (S. Bonassi)

Tipologia progetto: clinico-epidemiologica osservazionale

Area di interesse: prevenzione primaria/secondaria

Soggetti cofinanziatori: Commissione Europea; Ministero della Salute; Associazione Italiana per la Ricerca sul Cancro

Background

L'epidemiologia molecolare è nata dalla necessità di integrare i recenti grandi sviluppi della biologia molecolare con disegni di studio più rigorosi e metodiche di indagine più complesse tipici dell'epidemiologia tradizionale. Tra i campi che maggiormente hanno beneficiato di questo interscambio vi è l'analisi statistica dei dati biologici individuali (biomarcatori). Nei primi lavori l'analisi dati era insufficiente ed imprecisa mentre attualmente si è passati alla definizione di protocolli di analisi da applicarsi specificamente a ciascun biomarcatore. In tempi più recenti l'evoluzione di nuove tecniche di biologia molecolare - definite di high throughput - ha creato la necessità di sviluppare nuovi approcci informatici e statistici adatti all'elaborazione di grandi matrici di dati. In particolare, la tecnologia dei microarray permette di analizzare simultaneamente l'espressione di decine di migliaia di geni con un singolo esperimento fotografando il "livello di attività" dei geni in un individuo. Esaminando il profilo di espressione dei geni di un individuo affetto da una malattia o esposto ad un agente genotossico e comparando l'espressione genica in individui malati rispetto ai sani o esposti rispetto ai controlli, è possibile individuare nuovi geni la cui espressione è associata a determinati stati patologici o di esposizione e predire l'appartenenza di un individuo, per esempio, ad una specifica classe di malattie genetiche, unicamente in base al suo profilo di espressione.

Obiettivo generale del progetto ed eventuali obiettivi secondari

Nei tre anni oggetto del progetto lo scopo principale è quello di applicare metodiche di analisi statistica allo specifico settore dei biomarcatori utilizzando le tecniche che meglio si adattano alla particolare natura degli end-points osservati. Scopo finale è la definizione di standard metodologici per la conduzione di analisi statistiche in studi di popolazione su biomarcatori. A tal fine verranno utilizzati dati prodotti all'interno dell'Istituto o nell'ambito di collaborazioni nazionali ed internazionali. L'altro obiettivo - che riguarda la genomica - è quello di identificare, mediante tecniche statistiche applicate ai dati di espressione genica, i meccanismi di interazione tra geni che regolano i complessi processi di attivazione alla base dell'insorgenza di malattie. Questo obiettivo è perseguito, in particolare, mediante lo sviluppo di modelli matematici in grado di incorporare al loro interno il livello di interazione associato a tutte le potenziali "coalizioni" di geni, ed il successivo sviluppo di indici di rilevanza volti alla selezione dei geni più influenti. Tali modelli, basati sulla disciplina matematica nota come Teoria dei Giochi, sono da noi applicati allo studio di dati di microarray provenienti da soggetti principalmente in età pediatrica, come bambini affetti da neuroblastoma o bambini sani esposti ad inquinamento ambientale.

Impatto assistenziale certo o potenziale

Soprattutto nel campo oncologico la genomica ha aperto nuove prospettive per la classificazione molecolare delle patologie, la comprensione dell'eziologia dei tumori e l'individuazione di nuovi bersagli terapeutici.

Attività programmate 2009-2011 e risultati attesi

In rapporto alle caratteristiche di ciascun biomarcatore verrà messo a punto il modello statistico più adatto ad evidenziare eventuali relazioni con le caratteristiche degli individui. Sul versante della genomica verranno in particolare sviluppati i seguenti punti: disegno dello studio sperimentale; identificazione di metodologie volte alla "purificazione" e normalizzazione dei dati sperimentali; analisi mediante tecniche statistiche supervisionate e non-supervisionate. Risultato finale della ricerca sarà l'impostazione di linee guida per l'analisi statistica dei biomarcatori e dei microarray. Nel corso del primo anno si prevede di studiare nuovi modelli di gioco definiti su "network di co-

Programmazione 2009-2001

espressione", ovvero reti di geni basate su rapporti di correlazione tra coppie di geni in un set di dati di espressione genica. Tali metodologie di analisi statistica saranno sviluppate nell'ambito dei seguenti studi: identificazione di una signature genetica associata alla progressione di malattia nel neuroblastoma metastatico ad alto rischio; studio molecolare del microambiente midollare nei pazienti affetti da neuroblastoma; analisi dell'espressione genica di bambini affetti da medulloblastoma; analisi della frequenza di micronuclei in rapporto all'espressione genica di bambini esposti all'inquinamento atmosferico; analisi dell'espressione genica in coppie madri-figli in relazione all'esposizione attraverso la dieta.

Validazione di biomarcatori molecolari di esposizione, di effetto biologico precoce e di suscettibilità

Linea di ricerca: 1 – Prevenzione e Cancerogenesi

Programma: b - Biomarcatori biologici e molecolari di esposizione, di danno, di suscettibilità e di rischio di cancro

Responsabile scientifico: Marcello Ceppi

Altro personale della struttura partecipante al progetto: Cecilia Lando

Anno di inizio: 2009

Durata: 36 mesi

Parole chiave: epidemiologia molecolare; biomarker; polimorfismi genetici; danno al DNA; danno cromosomico

Altre strutture IST partecipanti: S.C. Epidemiologia e Biostatistica (D. Ugolini, R. Filiberti); S.S. Cancerogenesi Ambientale (C. Bognesi)

Altri Enti coinvolti: IARC, Lione, Francia (P. Boffetta); FIOH, Helsinki (H. Norppa); CSIRO, Adelaide, Australia (M. Fenech); Berkeley University, Berkeley, USA (N. Holland); CSPO, Firenze (M. Peluso); Università di Pisa (L. Migliore); National Yang Ming University Medical School, Taipei, Taiwan (W. Peter Chang); Istituto Superiore di Sanità, Roma (A. Zijno); Vrije Universiteit, Bruxelles, Belgio (M. Kirsh-Volders); IRCCS S. Raffaele Pisana, Roma (S. Bonassi)

Tipologia progetto: clinico-epidemiologica osservazionale

Area di interesse: prevenzione primaria/secondaria

Soggetti cofinanziatori: Agenzia Spaziale Italiana; Associazione Italiana per la Ricerca sul Cancro; Ministero della Difesa; Regione Emilia-Romagna; Compagnia di San Paolo; Fondazione CARIGE

Background

L'Epidemiologia Molecolare in campo oncologico è finalizzata a: i) conduzione di studi di tipo descrittivo e analitico per valutare il ruolo dell'interazione fra geni e ambiente nell'eziologia, nella storia naturale e nella prognosi delle malattie, ii) prevenzione della patologia e del danno genetico valutando i rischi a livello individuale e di popolazione ed identificando sottogruppi di individui suscettibili tramite screening genetico. In questi ultimi anni le ricerche sull'interazione fra geni e ambiente e sulla validazione dei biomarcatori predittivi di malattia sono state fra le più produttive in letteratura rendendo questa linea di ricerca tra le più promettenti nell'intento di comprendere il contributo dei fattori genetici, dei fattori di rischio ambientali e della loro interazione nell'eziologia del cancro.

Obiettivo generale del progetto ed eventuali obiettivi secondari

Il progetto si articola su diversi endpoints:

- Individuazione di indicatori precoci di cancerogenesi. Scopo della ricerca è di individuare quei biomarcatori che misurati in individui sani permettano di valutare precocemente il rischio di cancro e quindi di consentire tempestivi interventi di prevenzione primaria e/o di chemoprevenzione in gruppi esposti a sostanze sospette cancerogene. Lo studio della associazione fra questi marcatori e lo sviluppo di tumori contribuisce inoltre alla migliore comprensione delle fasi iniziali della cancerogenesi.
- Studio dell'interazione gene-ambiente. Molti geni implicati nella cancerogenesi sono polimorfici e le diverse varianti possono avere differente attività enzimatica con conseguente alterazione della risposta cellulare agli agenti genotossici. I polimorfismi possono pertanto essere correlati ad un diverso rischio di cancro e l'individuazione di sottopopolazioni suscettibili permetterebbe di programmare interventi di prevenzione. L'espressione e l'attività di diversi enzimi variano con l'età e questo può influire sulla sensibilità ai cancerogeni nel corso della vita.
- Validazione di biomarcatori genetici in studi di popolazioni umane. Uno dei maggiori limiti nell'utilizzo dei biomarcatori è che non sempre è chiaro ciò che essi effettivamente misurano. Nell'ottica di un uso clinico di questi indicatori diventa quindi fondamentale la loro validazione in termini di predittori di rischio di cancro. Seguire nel tempo anche retrospettivamente popolazioni umane sottoposte a test genetici fornirà importanti indicazioni per chiarire il ruolo dei biomarcatori nell'ambito della cancerogenesi.

Impatto assistenziale certo o potenziale

Fra le principali aspettative dell'utilizzo di marcatori genetici in clinica vi è la possibilità di perfezionare le procedure diagnostiche ed abbreviarne i tempi, la personalizzazione delle terapie in funzione del background genetico individuale, la comprensione anticipata dell'efficacia del trattamento e l'individuazione precoce (o la predizione) degli eventi avversi.

Programmazione 2009-2001

Attività programmate 2009-2011 e risultati attesi

I risultati ottenuti nel triennio 2006-8 sulle validazione del saggio del micronucleo nei linfociti ha permesso l'avvio di un nuovo progetto per la validazione del MN in saggi condotti sulla mucosa esfoliata della bocca, un test di facile attuazione in quanto meno invasivo del prelievo ematico. Il progetto denominato, HUMN-xI, è una naturale continuazione del progetto HUMN e prevede di completare entro il 2009 la raccolta dei dati. Successivamente si procederà alla standardizzazione dei dati provenienti dai diversi laboratori che partecipano al progetto. Un problema che si incontra nell'analisi di biomarcatori utilizzando fonti di dati differenti è l'eccesso di variabilità della misurazione dovuta a procedure di laboratorio non ancora completamente standardizzate, in particolare per i biomarcatori più innovativi. Allo scopo verranno attuate procedure di standardizzazione statistica e/o si farà ricorso a modelli statistici adatti all'analisi di dati organizzati in "cluster" (ad es. un laboratorio), che tengano conto della variabilità "dentro cluster" rispetto a quella "tra cluster". Le alterazioni del materiale genico verranno messe in relazione allo stato di malattia degli individui e/o all'esposizione a sostanze genotossiche rispetto ad individui sani/non esposti al fine di validare il biomarcatore in quanto indicatore di danno/esposizione a cancerogeni. Per quanto concerne lo studio dell'interazione gene-ambiente è stato avviato un progetto sul ruolo eziologico e patogenetico dei polimorfismi genetici in pazienti affetti da mesotelioma. In ambito allo studio promosso dal Ministero della Difesa sui soldati Italiani in Iraq si prevede di procedere all'analisi dei dati acquisiti e di coordinare il comitato che scriverà la relazione conclusiva. Infine, continua la collaborazione con le ARPA dell'Emilia-Romagna, per valutare la fattibilità di uno studio di biomonitoraggio in cittadini residenti in prossimità di inceneritori di rifiuti urbani.