

# DNA-microarray: nuovi approcci sperimentali per lo studio dei gemelli

Simonetta Pulciani<sup>(a)</sup>, Anna Di Lonardo<sup>(b)</sup>, Corrado Fagnani<sup>(a)</sup>,  
Cristina D'Ippolito<sup>(a)</sup>, Roberto Bompreszi<sup>(c)</sup> e Maria Antonietta Stazi<sup>(a)</sup>

<sup>(a)</sup>Centro Nazionale di Epidemiologia, Sorveglianza e Promozione della Salute,

<sup>(b)</sup>Dipartimento di Malattie Infettive, Parassitarie ed Immunomediate, Istituto Superiore di Sanità, Roma

<sup>(c)</sup>Barrow Neurological Institute, St. Joseph's Hospital and Medical Center, Phoenix (AZ), USA

**Riassunto.** Il completamento del Progetto Genoma Umano e le innovazioni introdotte in biotecnologia stanno cambiando il modo di studiare i gemelli. Alcuni studi molecolari su popolazioni di gemelli monozigoti (MZ) discordanti con applicazione di microarray sono qui riportati. I microarray consistono in una disposizione ordinata di un numero elevato di sonde (DNA, RNA o proteine), su un supporto solido. La tecnologia del microarray permette un'analisi globale dell'espressione genica e quindi potrebbe definire i meccanismi molecolari della discordanza nei gemelli MZ, ad esempio i meccanismi epigenetici. I dati derivanti dall'applicazione dei microarray allo studio dei gemelli consentiranno, tramite l'ausilio della bioinformatica, di comprendere il ruolo dei fattori genetici ed ambientali nello sviluppo di patologie umane, con conseguenti implicazioni diagnostiche e terapeutiche.

*Parole chiave:* microarray analysis, gemelli, modificazioni epigenetiche, espressione genica.

**Summary** (*DNA-microarray: new technological approaches on twin studies*). The completion of the Human Genome Project, and the innovations introduced in biotechnology are changing how to study twins. Here, we summarize some molecular studies performed on populations of discordant monozygotic twins (MZ) applying microarrays. Microarrays are an orderly arrangement of high numbers of probes (DNA, RNA or proteins), immobilized onto a matrix. The microarray approach allows a global analysis of gene expression, and therefore might point out the molecular mechanisms of MZ twins' discordance, such as epigenetic mechanisms. The application of microarray to twin studies will help better define, through bioinformatics, the role of genes and environment in the development of human diseases, thus suggesting new diagnostic and therapeutic approaches.

*Key words:* microarray analysis, twins, epigenetic modifications, gene expression.

## INTRODUZIONE

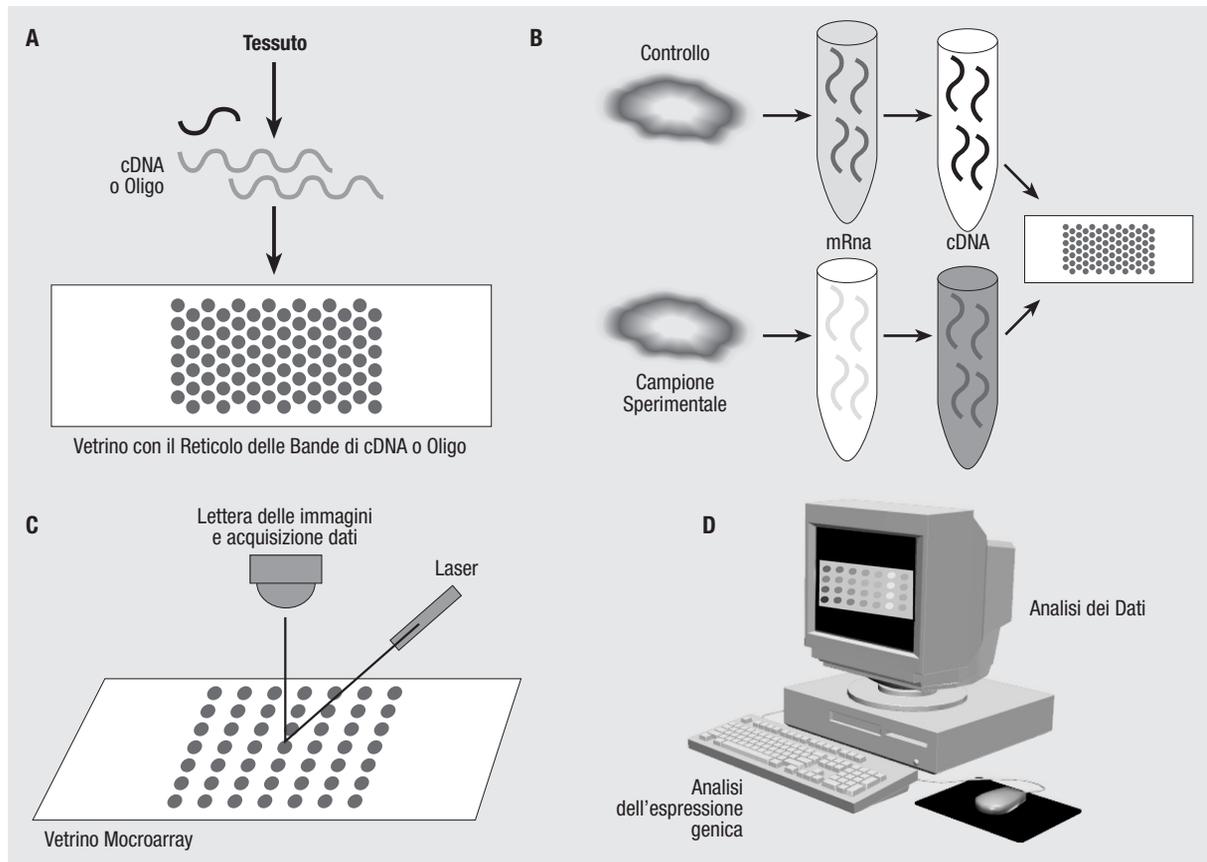
Dal punto di vista scientifico, i gemelli hanno rappresentato non solo un curioso fenomeno naturale ma anche un'importante fonte di studio. Il loro impiego quale modello per l'analisi dei fattori genetici e ambientali coinvolti nello sviluppo di un determinato fenotipo, normale o patologico, fu postulato per la prima volta da Galton nel 1874, nel suo libro *English men of science*, come descritto in Boomsma *et al.* [1]. Pur non essendo allora ancora nota la differenza tra gemelli monozigoti (MZ) e dizigoti (DZ), l'intuizione dell'utilità dei gemelli diede avvio a nuovi studi di genetica che avrebbero in seguito condotto alla realizzazione dei Registri Gemelli (RG) [2, 3].

Nei RG sono riportate informazioni relative a caratteristiche fenotipiche, abitudini di vita ed esposizioni a rischi ambientali [4-6]. Tuttavia la complessità di dati contenuta in tali registri è rimasta fino a poco tempo fa ad uso quasi esclusivo dell'epidemiologia genetica [7]. Più recentemente, le innovazioni tecnologiche e le

conseguenti applicazioni in campo biomedico hanno portato al completamento di vari Progetti Genoma, che, nel contesto del continuo processo d'acquisizione di conoscenza, hanno aperto nuovi orizzonti nella prospettiva della cura delle malattie. La genetica, muovendosi al di là dell'epidemiologia, può ora contare su supporti sperimentali in grado di fornire risposte sempre più precise a patologie complesse quali il cancro, il diabete, le malattie cardiovascolari, l'artrite, ecc. [8-10], e l'esistenza dei RG rappresenta la chiave di volta per l'applicazione di nuove tecnologie.

Una delle nuove tecniche più promettenti per ottenere risposte in campo biomedico è quella dei DNA-microarray [11].

In questa rassegna ci proponiamo di illustrare i principi fondamentali di tale tecnica, le possibili applicazioni relativamente all'epidemiologia genetica, con particolare attenzione allo studio dei gemelli, le problematiche derivanti dal suo utilizzo e le prospettive future.



**Fig. 1** | Tecnica dei DNA microarray.

Le sonde di DNA derivano da sequenze geniche clonate da vari tessuti e sequenziate. Esse possono essere cDNA amplificati per PCR o oligonucleotidi di sequenza nota, e vengono depositate sui vetrini a formare il reticolo del microarray mediante sistemi automatici (1A). Gli RNA estratti dal campione e dal controllo vengono convertiti in cDNA, e marcati rispettivamente con il fluoroforo Cy3 e Cy5. I due cDNA, così ottenuti, vengono entrambi ibridati su uno stesso vetrino di DNA-microarray (1B).

Dopo il processo di ibridazione il vetrino viene lavato per eliminare sequenze non perfettamente appaiate. Successivamente viene letto con degli strumenti (Scanner) capaci di rilevare i segnali dei due fluorofori ed acquisire i dati della loro intensità (1C).

I dati ottenuti vengono analizzati ed elaborati con programmi informatici appropriati per determinarne il significato biologico (1D).

## LA TECNICA DEI DNA-MICROARRAY

La tecnica dei DNA-microarray deriva dai protocolli di Southern e Northern blot che furono introdotti circa trent'anni fa come metodiche per l'analisi di sequenze di DNA e RNA, rispettivamente [11, 12]. Essa è stata resa possibile dai dati ottenuti dal sequenziamento del DNA nei vari Progetti Genoma. Il principio, comune a queste metodiche, è elementare e si basa sull'ibridazione di sequenze di DNA complementari con sonde radioattive o chemoluminescenti a sequenza nota. La caratteristica distintiva dei microarray è la possibilità dell'analisi simultanea di migliaia di geni in un singolo esperimento.

In Figura 1 è schematizzata la procedura sperimentale adottata nell'analisi dell'espressione genica che rappresenta la problematica più frequente per cui i DNA-microarray sono applicati.

Le metodiche che consentono la preparazione dei microarray possono essere suddivise in due categorie (Figura 2). Una categoria utilizza sonde sintetizzate *in vitro*, frammenti di cDNA ottenuti da tessuti e sottoposti ad

amplificazione genica (PCR) oppure oligonucleotidi. Tali sonde vengono successivamente depositate sul vetrino [13] (Figura 2A). L'altra categoria utilizza tecnologie che permettono di sintetizzare sonde di oligonucleotidi direttamente sul supporto solido (Figura 2B e 2C).

La prima tecnologia ideata per la sintesi diretta d'oligonucleotidi *in situ*, mediante reazioni fotochimiche, è nata adattando le tecnologie dell'industria dei semiconduttori alla biologia molecolare, ed è stata sviluppata dalla società Affymetrix. I vetrini così preparati prendono il nome di GeneChips [14,15] (Figura 2B). Inoltre, più recentemente è stato introdotto un metodo alternativo che prevede la deposizione, sempre ad alta densità su vetrino pretrattato, d'oligonucleotidi, anch'essi pre-sintetizzati *in vitro*, attraverso la tecnica cosiddetta a getto o Ink-jet [16] (Figura 2C).

In questa rassegna non tratteremo dei sistemi Affymetrix e Ink-jet. In particolare per i GeneChips, i costi elevati, i relativi brevetti e la complessa strumentazione ne rendono la realizzazione perseguibile solo dall'industria. Pertanto focalizzeremo la nostra attenzione sulle tecnologie in cui i

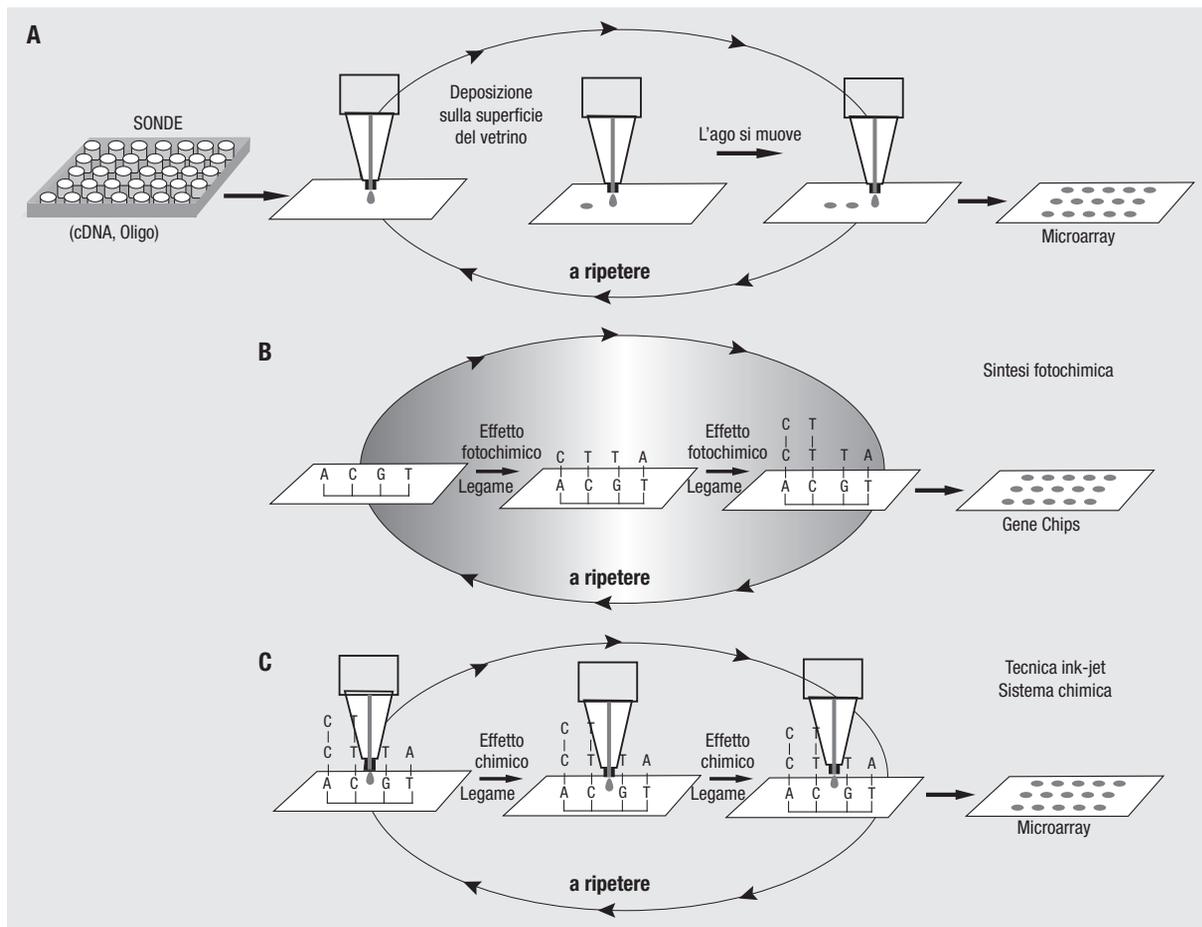
vetrini sono realizzati per deposizione di sonde preparate *in vitro*.

La tecnica dei DNA-microarray è stata ideata da Pat Brown, che, con l'intento di analizzare il coordinamento dell'espressione genica in cellule di lievito durante le varie fasi cellulari, ne ha prodotto la prima pubblicazione scientifica nel 1995 [13]. Nella sua concezione iniziale la metodica si affidava ad un sistema automatico in grado di deporre in maniera ben ordinata e riproducibile un alto numero di specifiche sonde di cDNA (ottenute per amplificazione per PCR) su un vetrino chimicamente trattato. Si otteneva così un micro-schieramento di sonde, da cui il nome microarray, su cui il campione da valutare s'ibridizzava in maniera proporzionale alla complementarità delle sequenze in esso contenute. La tecnica dei microarray è rimasta invariata nel suo principio base, ma considerevoli innovazioni sono state apportate agli strumenti e a tutto il sistema di produzione dei DNA-microarray. In linea generale, i dispositivi di

dispensazione delle sonde, Arrayer o Printer o Spotter [11, 17], sono fondamentalmente dei robot controllati da computer. Il braccio meccanico del robot dispone di speciali testine che raccolgono le sonde dal piatto dove sono conservate e, con precisione micrometrica, le depositano sul vetrino nella posizione e nell'ordine stabiliti, tramite diretto contatto con la superficie del vetrino stesso.

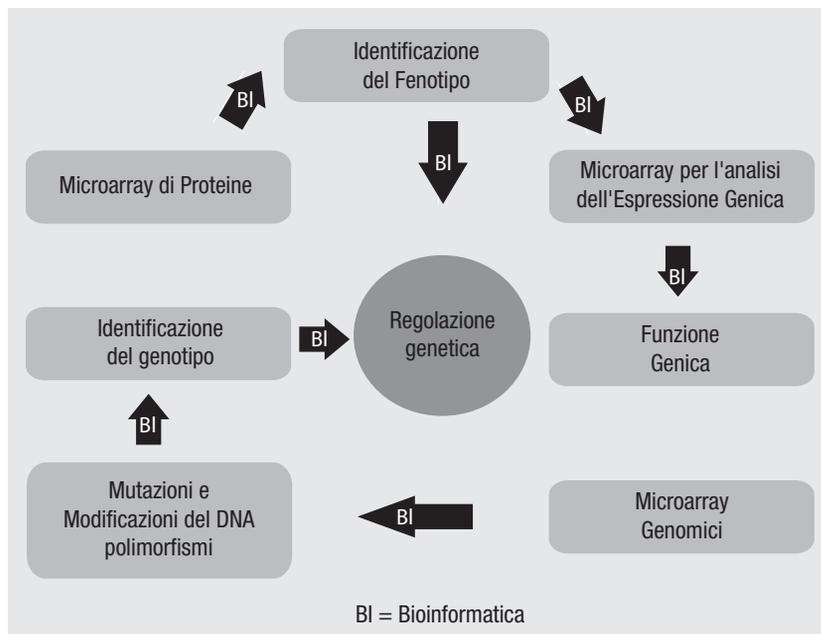
Inizialmente le sonde erano costituite da frammenti di cDNA amplificati per PCR, di lunghezze non superiori a 2000 basi, mentre gli oligonucleotidi da 50-70 basi pre-sintetizzati *in vitro* costituiscono la versione più recente dei microarray.

Una volta depositata sul supporto, la sonda copre un'area (spot) del diametro compreso tra 80 e 150 micron (*Figura 1A*); l'intervallo tra uno spot e il successivo è di circa 100 micron e generalmente ogni vetrino contiene un reticolo di almeno 1000 spot, mentre con i più moderni microarray si raggiungono facilmente i 30 000 spot per vetrino. In alcuni casi gli spot possono



**Fig. 2** | Tecnologie per la produzione dei vetrini microarray.

Le sonde prodotte *in vitro* vengono aspirate da microaghi e deposte per contatto o "spruzzo" su vetrini chimicamente trattati. L'ago si muove e il processo si ripete più volte fino a formare il microreticolo (2A). Nella sintesi fotochimica *in situ* le sonde sono rappresentate esclusivamente da oligonucleotidi. Essi vengono sintetizzati per aggiunta progressiva di basi che vengono legate covalentemente per effetto fotochimico. Il processo si ripete più volte fino a formare il microreticolo (gene chips) (2B). Anche nella sintesi chimica *in situ* le sonde sono rappresentate esclusivamente da oligonucleotidi. In questo caso i reagenti, comprese le basi, vengono depositati progressivamente a "spruzzo" (tecnica ink-jet). Le basi vengono legate chimicamente, lo spruzzatore si muove e il processo si ripete più volte fino a formare il microreticolo (2C).



**Fig. 3** | Applicazioni della bioinformatica nei microarray. I risultati dei microarray per l'espressione genica, dei microarray genomici e di proteine vengono elaborati dalla bioinformatica. Si possono così ottenere dati sulla funzione genica, sul genotipo e sul fenotipo. L'insieme di queste informazioni permette di comprendere i meccanismi di regolazione inter- e intra-genica.

avere dimensioni maggiori di 250 micron e allora si parla di DNA-microarray.

Per quanto riguarda la preparazione del campione da analizzare sui vetrini, particolare attenzione va posta alle procedure d'estrazione e preservazione degli RNA per i quali sono stati sviluppati specifici protocolli. Gli RNA messenger del campione e del controllo sono retro-trascritti in cDNA e, nel corso di tale reazione, due distinti fluorocromi, tipicamente le cianine Cy3 e Cy5, sono incorporati per funzionare da marcatori dei campioni di partenza (Figura 1B). Dopodiché i campioni possono essere mescolati e posti ad ibridizzare insieme su un unico vetrino da microarray.

I fluorofori Cy3 e Cy5 emettono segnali fluorescenti a differente lunghezza d'onda permettendo così di identificare il campione e il controllo. Il livello d'espressione genica è calcolato in base all'intensità del segnale, registrata da ciascuno spot presente sul vetrino e rappresentativo dei vari geni. La possibilità teorica della competizione tra i due campioni per le sonde è certamente trascurabile, dato il grande eccesso di molecole delle sonde presenti sul vetrino, per cui i campioni s'ibridizzano in maniera proporzionale alla quantità di trascritto presente in origine.

Dopo il processo d'ibridazione, il vetrino è sottoposto ad opportuni lavaggi per eliminare il campione non ibridizzato, ed è esaminato tramite uno strumento (scanner) provvisto di due fasci di luce laser, a lunghezza d'onda specifica, in grado di eccitare i due fluorofori (Figura 1C) il cui segnale fluorescente di diversa lunghezza d'onda è raccolto da rilevatori in due canali separati.

L'intensità del segnale ottenuta dall'intero vetrino è sottratta ai valori di fondo, e viene calcolato il rapporto tra i due canali. I valori del rapporto sono poi convertiti in immagine dove a ciascuno spot corrisponde un'intensità di una scala di pseudo-colori con gradazioni del

rosso, che in genere è usato per indicare l'aumentato livello d'espressione nel campione rispetto al controllo. Viceversa, valori del rapporto negativi, corrispondenti a livelli d'espressione più alti nel controllo relativamente al campione, sono indicati da varie gradazioni del verde. Il colore giallo è usato per valori del rapporto vicini all'unità. Ne risulta così un'immagine combinata che riproduce il reticolo del DNA-microarray e consente una visione immediata degli spot con intensità di segnale più elevata in un canale che nell'altro, corrispondente in definitiva al diverso livello di contenuto di RNA messenger nei campioni ibridizzati.

L'insieme dei dati è poi sottoposto ad analisi nel contesto di un singolo esperimento, e/o d'esperimenti multipli atti a identificare variazioni biologiche (Figura 1D). L'analisi statistica è una parte critica ed essenziale della metodica. Infatti, i dati che si ottengono con i DNA-microarray sono innumerevoli, e l'elaborazione attraverso sistemi informatici, che rappresenta l'ultima fase della sperimentazione, è delicata e complessa [18, 19].

La tecnologia dei DNA-microarray, in circa un decennio dalla sua nascita, si è evoluta continuamente, ed oggi è possibile depositare sui vetrini anche sonde costituite da DNA genomico. Queste ultime costituiscono la base per la preparazione dei microarray genomici, utilizzati per l'analisi d'amplificazioni e delezioni geniche o mutazioni epigenetiche [20-22].

L'idea del microarray, cioè di organizzare su uno stesso vetrino un cospicuo numero di sonde, è stata estesa allo studio di proteine e tessuti, rendendo così oggi disponibili microarray per tali materiali biologici [23, 24].

I risultati acquisiti dai diversi tipi di microarray (DNA, genomici, proteine, ecc.) necessitano dell'ausilio di sistemi informatici per essere analizzati [25, 26], e la bioinformatica, scienza nata dall'esigenza di catalogare ed elaborare i dati acquisiti attraverso i Progetti Genoma,

trova una delle sue più adatte applicazioni nel campo dei microarray. Con sempre maggiori sofisticazioni c'è la speranza di giungere alla descrizione dei più complessi fenomeni biologici.

In *Figura 3* sono illustrati i principali impieghi della bioinformatica nel campo della biologia.

Ad oggi, sono stati implementati numerosi sistemi operativi per analizzare e catalogare i dati di un esperimento singolo e per confrontare ed integrare quelli di esperimenti distinti [25-27]. La ricerca biologica è in continua evoluzione e la bioinformatica avanza di pari passo [25].

Per un approfondimento maggiore sulla tecnologia dei microarray e possibili protocolli si possono consultare specifici siti Internet (*Tabella 1*).

### APPLICAZIONI DELLA TECNICA DEI MICROARRAY

Prima di descrivere le applicazioni dei microarray alla ricerca biomedica, ci sembra opportuno fare un breve cenno sui progetti di ricerca, avviati agli inizi degli anni '80, e volti al sequenziamento del patrimonio genetico di differenti organismi, con particolare riferimento all'uomo [28].

Il Progetto Genoma Umano (PGU) è uno studio collaborativo a livello internazionale i cui risultati sono stati ottenuti attraverso sforzi congiunti e resi disponibili all'intera comunità scientifica in tempo reale; la sua prima fase è stata completata nel 2001, a circa cinquant'anni dalla scoperta del DNA [8, 9, 28]. Tale progetto costituisce un'instimabile risorsa: la conoscenza dell'intera sequenza delle basi azotate che formano il nostro patrimonio genetico. L'archiviazione delle sequenze di tutti i cromosomi può sembrare, a prima vista, di scarso significato scientifico, ma rappresenta la *condicio sine qua non* per definire la funzione, l'organizzazione, la regolazione di ognuna di esse; è in definitiva la lettura del codice della vita umana [25, 28, 29].

In parallelo al PGU, si sono sviluppati e realizzati innumerevoli progetti, focalizzati sul sequenziamento e alla caratterizzazione strutturale e funzionale del genoma d'altri esseri viventi [28]. Nel complesso di un avanzamento globale di conoscenza, tali studi forniranno un contributo essenziale alle scienze biomediche, permet-

tendo di delineare i meccanismi molecolari di patologie semplici e complesse, i loro fondamenti genetici ed il ruolo dei fattori ambientali [27-29]. L'obiettivo però è ancora lontano: giungere alla comprensione delle funzioni dei geni e dei loro meccanismi di regolazione [30-32]. I risultati ottenuti finora necessitano di un lavoro di rifinitura tramite la verifica delle sequenze geniche già ottenute, l'allineamento delle sequenze di diverse specie biologiche e le acquisizioni dei numerosi polimorfismi, che sono alla base della variabilità della specie umana.

L'insieme delle ricerche dedite al perseguimento di tali obiettivi costituisce una nuova era della biologia che è stata definita era Post-Genomica [28-32]. L'era Post-Genomica presuppone una "visione e versione globale" della sperimentazione perché l'analisi non è limitata a singoli geni, ma interessa l'intero genoma e pertanto le tecniche e i protocolli della biologia molecolare classica sono obsoleti. Con l'avvento dell'era Post-Genomica sono stati introdotti nuovi termini, quali Transcriptoma e Proteoma, che descrivono, rispettivamente, le diverse molecole di RNA e l'insieme delle proteine presenti in una cellula o organismo in un preciso istante. Analogamente, le scienze che studiano questi componenti cellulari e i loro processi di trascrizione e traduzione sono chiamate Transcriptomica e Proteomica [28-32].

La ricerca Post-Genomica deve far affidamento su tecnologie in grado di esaminare innumerevoli campioni simultaneamente, ed i microarray rappresentano uno degli approcci sperimentali più appropriati a tale scopo [27, 30-32]. La tecnica dei DNA-microarray, nata per l'analisi dell'espressione genica differenziale in cellule di lievito [13], nell'era Post-Genomica ha subito evoluzioni che le hanno consentito di adattarsi allo studio dell'organizzazione strutturale e funzionale del genoma umano [27, 33]. Di conseguenza, le potenzialità della tecnica dei DNA-microarray si sono ampliate ed estese dall'analisi dell'espressione genica alla caratterizzazione d'amplificazioni e delezioni geniche, allo studio dei siti di legame dei fattori di trascrizione e all'analisi di mutazioni epigenetiche [22, 30]. Inoltre, mediante i Gene-chips è possibile sia verificare sequenze già presenti in banche dati che individuare la presenza di polimorfismi e quindi d'eventuali mutazioni puntiformi [31, 32].

**Tabella 1** | Alcuni siti web dedicati ai microarray

Istituzione	Indirizzo web
National Human Genome Research Institute - NIH - USA	<a href="http://research.nhgri.nih.gov">http://research.nhgri.nih.gov</a>
European Bioinformatic Institute - UK	<a href="http://www.ebi.ac.uk/microarray">http://www.ebi.ac.uk/microarray</a>
National Institute of Environmental Health Sciences - NIH - USA	<a href="http://dir.niehs.nih.gov/microarray">http://dir.niehs.nih.gov/microarray</a>
University of California - San Francisco - USA	<a href="http://derisilab.ucsf.edu/microarray/index.html">http://derisilab.ucsf.edu/microarray/index.html</a>
European Science Foundation - UK	<a href="http://www.functionalgenomics.org.uk">http://www.functionalgenomics.org.uk</a>
CRIBI - Università di Padova - Italia	<a href="http://microcribi.cribi.unipd.it/i-index.htm">http://microcribi.cribi.unipd.it/i-index.htm</a>
Fondazione "San Raffaele" - Milano - Italia	<a href="http://www.sanraffaele.org">http://www.sanraffaele.org</a>

Indirizzi web - data ultima consultazione: 16 maggio 2006.

Tra le applicazioni più comuni dei DNA-microarray vi sono l'analisi dell'espressione genica differenziale tra mutanti cellulari e linee parentali, o la comparazione dei profili in differenti condizioni di crescita (ad esempio specifici elementi nutritivi, tossici, farmaci, ecc.). I risultati, ottenuti in simili esperimenti, sono di fondamentale importanza al fine di identificare i meccanismi di regolazione genica e le possibili vie biosintetiche [34]. Tale approccio è stato poi applicato all'analisi genetica per lo studio del cancro, dei processi infettivi e d'invecchiamento [35-45], e i dati derivanti da questi studi ne hanno dimostrato la validità per l'analisi molecolare a largo spettro di vari processi patologici.

L'utilizzo dei DNA-microarray per lo studio delle neoplasie è emblematico delle sue grandi potenzialità. Mentre, da sempre, la diagnosi oncologica si è affidata alla valutazione clinica ed istologica, oggi i DNA-microarray contribuiscono in modo determinante alla diagnosi delle patologie tumorali. Ne permettono un'analisi molecolare attraverso l'espressione genica che ne definisce le alterazioni strutturali cromosomiche, quali delezioni o amplificazioni geniche [35, 36, 40-44], o più in generale ne permettono la caratterizzazione molecolo-funzionale che è in grado di determinare una diagnosi differenziale accurata tra lesioni maligne e premaligne. È stato osservato che l'analisi del profilo d'espressione di tumori del sistema emopoietico permette di stabilire l'esatto tipo cellulare cui il tumore appartiene, condizione necessaria per formulare terapie efficaci. Golub *et al.* hanno studiato gli RNA messaggeri, isolati da cellule del midollo osseo di pazienti affetti da leucemia mieloide acuta (AML) o leucemia linfocitica acuta (ALL) [42]. Dopo ibridazione su vetrini di microarray, rappresentativi di circa 6800 geni umani, e la relativa analisi statistica, una selezione di 50 geni differenzialmente espressi si è dimostrata capace di identificare i due tipi di leucemia [42]. Così come un patologo può riconoscere ALL da AML osservando al microscopio i vetrini, altrettanto bene, e forse meglio, i microarray permettono di distinguere i due tipi di neoplasie.

Analogamente, Brown *et al.*, esaminando linfomi del tipo non-Hodgkin con vetrini di microarray (con circa 15 000 geni umani), hanno rilevato differenze sostanziali nell'espressione genica tra i tumori rispondenti e non alle terapie [43].

La tecnica dei microarray ha fornito contributi notevoli anche alla classificazione molecolare di melanomi e tumori della ghiandola mammaria, non solo mediante l'analisi dell'espressione genica, ma anche attraverso l'analisi delle delezioni e amplificazioni geniche con la tecnica definita Comparative Genomic Hybridization microarray (CGH-microarray) [40, 44].

Il protocollo per il CGH-microarray deriva dalla combinazione dei microarray con la tecnica *Comparative Genomic Hybridization*, quest'ultima concepita proprio per analizzare amplificazioni e delezioni geniche [44]. I vantaggi del CGH-microarray consistono nella possibilità non solo di esaminare uno stesso campione con un elevato numero di sonde, ma anche di offrire una migliore risoluzione nell'analisi delle copie di geni.

La tecnologia dei DNA-microarray è utilizzata anche

nel campo delle patologie infettive per analizzare i profili d'espressione sia dell'ospite sia del patogeno [46]. I patogeni maggiormente studiati sono rappresentati da batteri e miceti, i cui genomi sono stati sequenziati, completamente o in larga misura (1000 geni per batterio e 5-6000 per micete) [37, 38]. I risvolti pratici di tale approccio sono considerevoli, giacché è possibile così definire le basi genetiche della patogenicità e sviluppare nuove strategie preventive e terapeutiche [46].

Infine, in un campo differente ma di altrettanto alto interesse, i microarray hanno fornito evidenza di alterazione d'espressione di determinati geni in correlazione con un ritardo dei processi d'invecchiamento [39, 47, 48]. In particolare, studi su modelli murini hanno sottolineato il ruolo di geni espressi in situazioni di stress quali HIC-5 (un fattore trascrizionale indotto da processi d'ossidazione cellulare), ATP sintetasi, gene ERV1 (coinvolto con le funzioni del DNA mitocondriale) ed altri ancora [47, 48].

Nel complesso i dati derivati dall'analisi di microarray danno risultati sui determinanti molecolari delle più svariate patologie che si traducono in ultima analisi in acquisizione d'informazioni sulla funzione dei geni, dei loro meccanismi regolatori e delle vie biosintetiche [27]. Proprio la conoscenza di questi ultimi è, infatti, il requisito essenziale per il raggiungimento dell'obiettivo centrale della ricerca Post-Genomica: ottenere nell'immediato futuro una caratterizzazione diagnostica accurata delle patologie umane.

Come discuteremo nella sezione successiva, l'epidemiologia genetica e la gemellologia non saranno esenti da questa rivoluzione di metodologie.

## LA TECNICA DEI MICROARRAY NELLO STUDIO DEI GEMELLI

L'epidemiologia ha come oggetto l'osservazione d'individui con o senza patologie, ed è stata da sempre una scienza basata sull'analisi degli eventi naturali. Proprio il confronto tra le categorie d'individui sani e malati, ed in particolare il confronto tra le frequenze di esposizione a determinati fattori in tali categorie, ha permesso di stimare la frequenza delle varie patologie nella popolazione e anche l'associazione tra queste ed i fattori di interesse, per poi suggerire misure di prevenzione [49].

L'epidemiologia genetica, in particolare, ha l'obiettivo di definire come i fattori genetici e ambientali interagiscano nello sviluppo delle malattie complesse [7, 49], e i gemelli rappresentano uno dei possibili modelli di studio. L'incidenza delle nascite gemellari varia nelle diverse popolazioni, con una media di un parto gemellare ogni cento.

A seconda dell'origine da uno o più oociti, i gemelli possono essere classificati in monozigoti (MZ) e dizigoti (DZ) [1]. I gemelli MZ, che rappresentano 1/3 del totale, condividono lo stesso patrimonio genetico ed ogni variazione fenotipica fra di loro può quindi essere attribuita o a fattori ambientali non condivisi o a fenomeni epigenetici/post-trascrizionali. I gemelli DZ, invece, hanno in comune solo la metà dei loro geni, come dei normali fratelli. È plausibile ipotizzare

che i gemelli DZ abbiano una maggiore condivisione d'esposizione a fattori ambientali rispetto a fratelli nati a distanza di tempo, ma, in ogni caso, la variazione fenotipica fra loro può essere attribuita sia a fattori ambientali sia a fattori genetici [1-6]. Il confronto tra le concordanze o le correlazioni osservate per un certo carattere nei MZ e nei DZ consente di stabilire il ruolo dei geni e dell'ambiente nell'espressione del fenotipo; in particolare, esso permette di stimare l'ereditabilità del carattere, cioè la proporzione della variabilità fenotipica inter-individuale dovuta ai fattori genetici.

Il potenziale degli studi sui gemelli è aumentato notevolmente in seguito alla nascita dei Registri Gemelli (RG) su base di popolazione. Questi rappresentano delle risorse inestimabili, poiché contengono un'enorme quantità di dati sul fenotipo, abitudini di vita, malattie ed esposizione a rischi ambientali in un grande numero di gemelli e delle loro famiglie [1-6]. Un potente strumento di indagine si ottiene quando i dati provenienti da diversi RG vengono analizzati in maniera combinata, come nel caso del progetto europeo GenomEUtwin [2-3], che coinvolge otto RG, incluso quello italiano [2].

L'analisi dei gemelli può essere d'aiuto nello stabilire le influenze genetiche ed ambientali nella predisposizione alle malattie. Tuttavia, definire come gli stimoli ambientali influiscano sulle concordanze e/o discordanze fenotipiche dei gemelli è un obiettivo non semplice.

Finora i gemelli sono stati studiati principalmente usando i metodi epidemiologici classici; oggi i nuovi mezzi tecnologici, derivati dalla ricerca Genomica e Post-Genomica, offrono l'opportunità di analizzare e interpretare a livello molecolare i dati dei RG [1-10, 28].

I dati ottenuti dal PGU in associazione con quelli estrapolati dai RG permetteranno di definire l'influenza dell'ambiente e/o del genotipo nell'espressione di un particolare fenotipo [7, 49].

Finora con gli approcci epidemiologici classici e le tecniche di clonaggio molecolare sono stati caratterizzati soltanto geni associati a malattie di tipo mendeliano, che coinvolgono cioè un solo gene [50]. In futuro la maggiore conoscenza del genoma umano e dei meccanismi di regolazione genica permetteranno l'analisi genetico-molecolare di patologie complesse, che risultano dall'interazione di due o più geni [51]. In questo contesto i microarray rappresentano una delle tecnologie più promettenti nel campo della ricerca dedicata all'analisi del genoma umano e delle malattie complesse [52].

Scopo principale di questo paragrafo è descrivere le possibili opportunità offerte dalla tecnologia dei microarray nello studio dei gemelli; di seguito sono presentati due esempi di studi volti alla caratterizzazione dei profili d'espressione in gemelli MZ discordanti per patologia.

Un gruppo di ricercatori ha recentemente studiato il profilo d'espressione di plasmacellule CD138+ e CD138- isolate dal midollo osseo di una coppia di gemelli MZ discordanti per mieloma multiplo (MM) [53]. Gli esperimenti sono stati condotti utilizzando piattaforme di microarray (tipo HG-U95av2) della ditta Affymetrix che contengono più di 12 000 geni umani. L'analisi degli RNA ha messo in evidenza, nelle cellule CD138+ del gemello affetto da MM, la

sovraespressione di duecentonovantasei geni e la sottoespressione di centotré geni. Tra i geni sovraespressi sono stati individuati geni anti-apoptotici (FLIP, dad 1, mcl-1), oncogeni e fattori trascrizionali (Jun D, Xbp-1, FGFR-3) e geni collegati con l'attività ribosomale. Tra i geni sottoespressi, invece, sono stati rilevati principalmente quelli ad attività pro-apoptotica e di riparo e ricombinazione del DNA (RAD51). L'insieme di questi risultati ha messo in risalto un'aumentata attività metabolica e trascrizionale esclusivamente nelle cellule CD138+ provenienti dal gemello affetto da MM. Questi dati rispecchiano ciò che è noto avvenire nella cellula neoplastica e sottolineano il valore dell'approccio molecolare applicato ai gemelli.

L'altra patologia, presa in esame nei gemelli MZ discordanti mediante la tecnica dei microarray, è la malattia bipolare la cui eziologia rimane oscura, ma la cui possibile componente genetica è stata prospettata. Questa malattia psichiatrica è caratterizzata da un'alternanza di fasi depressive e fasi di eccitamento. In uno studio di KaKiuchi *et al.*, sono state arruolate due coppie di gemelli MZ discordanti per disturbi bipolari ed una coppia di gemelli sani, ed è stata analizzata l'espressione genica differenziale in cellule linfoblastoidi [54]. Tali esperimenti hanno evidenziato una ridotta espressione d'alcuni geni collegati al funzionamento del reticolo endoplasmatico, tra cui i geni XBP1 e HSPA5. Questi risultati sono di particolare interesse poiché l'espressione del gene HSPA5 è indotta dal composto chimico valproato che è utilizzato come stabilizzatore dell'umore ed è regolato dal gene XBP1. Inoltre il gene HSPA5 è localizzato sul cromosoma 22 in posizione 22q12, sito già precedentemente messo in relazione con le malattie di tipo bipolare [55].

I microarray di espressione sono stati utilizzati anche da ricercatori danesi ed indiani per confrontare coppie di gemelli di diverse età; tale approccio sperimentale ha permesso di rilevare una maggiore correlazione tra l'espressione genica nei gemelli MZ rispetto ai DZ [60, 61].

Se da una parte questi studi forniscono risposte, da un'altra sollevano nuovi interrogativi quali: "perché gemelli MZ pur condividendo lo stesso patrimonio genetico risultano discordanti per patologie e/o altri caratteri fenotipici?" Intuitivamente si è portati a pensare alla possibilità di mutazioni genetiche sopraggiunte in uno solo dei gemelli dopo la nascita. Secondo un'ipotesi prospettata più recentemente vi sarebbero meccanismi epigenetici (modificazioni epigenetiche) per i quali la struttura del DNA, ma non la sua sequenza, è modificata in maniera reversibile [56-58]. L'assetto sterico della doppia elica cambia favorendo o limitando la funzione degli enzimi deputati alla trascrizione del messaggio genetico. Tali fenomeni coinvolgono la struttura della cromatina ed alterano il processo di trascrizione, prima tappa nella sintesi proteica; le variazioni nel contenuto proteico possono alla fine dar luogo a differenti fenotipi. Tra le modificazioni epigenetiche meglio studiate vi sono la metilazione della citosina e l'acetilazione degli istoni [22, 56-58]. Recentemente il disegno architettonico dei microarray è stato sfruttato per costruire dei vetri atti a rilevare tali modificazioni su DNA genomico

a larga scala [21, 22]. Questi vetrini sono stati fino ad oggi impiegati soprattutto per lo studio dei meccanismi molecolari d'alcune forme tumorali [59]. Ci si auspica che in un futuro prossimo essi possano essere utilizzati anche nello studio dei gemelli e integrino dati ottenuti con metodi sperimentali classici riguardanti l'analisi d'alterazioni epigenetiche in coppie di gemelli MZ di differenti età [62]. A questo proposito, un gruppo spagnolo ha condotto delle ricerche dirette a caratterizzare sia la metilazione delle citosine che l'acetilazione degli istoni in linfociti e in altri tessuti, come ad esempio la mucosa boccale. I risultati di tali studi hanno messo in evidenza che le differenze epigenetiche tra gemelli MZ aumentano proporzionalmente non solo in rapporto all'età, ma anche all'ambiente non condiviso [62]. Lo stesso gruppo ha in seguito accertato, con esperimenti di microarray, che alle diversità delle mutazioni epigenetiche corrispondono differenze nell'espressione genica, fornendo così per la prima volta la prova sperimentale che l'ambiente contribuisce alle differenze fenotipiche osservate in gemelli MZ e verosimilmente attraverso meccanismi epigenetici.

Tuttavia, nel lavoro di KaKiuchi *et al.*, già descritto, e riguardante l'analisi molecolare in gemelli MZ discordanti per disturbi bipolari, i geni XBP1 e HSPA5, pur essendo modulati negativamente, non presentano alterazioni né genetiche né epigenetiche [54]. Queste osservazioni mettono in risalto che il cammino intrapreso nell'analisi delle mutazioni epigenetiche come possibili cause delle differenze fenotipiche tra gemelli MZ è valido, ma puntualizzano l'esigenza di più metodiche e approcci sperimentali al fine di determinare ulteriori meccanismi molecolari verosimilmente coinvolti [63]. Infatti, i lavori fino ad oggi pubblicati sui gemelli MZ discordanti per disturbi bipolari, compreso quello di Kakiuchi *et al.*, non sono riusciti a determinare le cause molecolari delle alterazioni, siano esse strutturali o epigenetiche, correlate ai processi di trascrizione. Inoltre, poiché il fenotipo è determinato dalle proteine biologicamente attive, l'analisi molecolare delle discordanze tra gemelli MZ non può prescindere dallo studio dei processi di traduzione e maturazione proteica. Infatti, gli studi sull'espressione genica in gemelli MZ discordanti forniscono un primo quadro delle proteine che potrebbero essere potenzialmente espresse, ma non determinano quali di queste giungeranno effettivamente a maturazione e quindi contribuiranno alla variabilità fenotipica. Dal punto di vista concettuale la risposta a tale quesito giunge dai microarray per le proteine. Tuttavia questo tipo di microarray trova un limite applicativo nelle difficoltà tecniche dovute alle caratteristiche chimico-fisiche delle proteine che mal si prestano a fenomeni d'ibridazione per complementarità [64], così ben riproducibili nel caso degli acidi nucleici.

In ogni caso ci si attende che gli avanzamenti tecnologici saranno in grado di superare gli ostacoli e le difficoltà tecniche ed in ultima analisi i microarray saranno in grado, attraverso sistemi operativi atti a gestirne l'elaborazione in modo organico ed interattivo [19, 25], di contribuire alla costruzione di mappe d'interazioni geniche [25, 27]. Tali mappe potranno costi-

tuire la base per determinare l'influenza dell'ambiente e della predisposizione genetica nei processi vitali e patologici.

## CONCLUSIONI

Le aspettative verso la ricerca Post-Genomica in generale, e dei microarray in particolare, sono molte; in sinergia con altre tecnologie derivate dai Progetti Genoma, quella dei microarray potrebbe dare risposte alle molteplici domande che le scienze biomediche si trovano quotidianamente ad affrontare [65].

Per poter comprendere i fenomeni patologici che interessano l'uomo, è indispensabile ed opportuno che l'uomo sia analizzato nella sua globalità, senza tralasciare l'influenza dell'ambiente, come sosteneva Ippocrate sin dall'antichità [10].

Innumerevoli progressi sono stati compiuti in questi ultimi anni, e un contributo rilevante può essere attribuito alle conoscenze acquisite con i Progetti Genoma. Tali conoscenze potranno, in un prossimo futuro, stabilire il ruolo della suscettibilità genetica e dei fattori ambientali nello sviluppo e nella progressione degli stati patologici [28]. L'identificazione di tutti i geni umani e la loro conseguente caratterizzazione, strutturale e funzionale, permetteranno di delineare una mappa genetica di patologie varie e produrre un cambiamento nel modo di formulare diagnosi e prognosi [10, 28]. Nuove tecnologie impensabili nel passato sono oggi disponibili e concorrono a tali obiettivi [28]. Una delle tecnologie più innovative è senz'altro quella dei microarray, sviluppatasi proprio grazie ai risultati conseguiti con i Progetti Genoma. Essa consente l'analisi simultanea e comparativa di migliaia di geni e dei loro prodotti in tessuti normali e patologici. Nell'ultimo decennio, grazie all'avvento di questa nuova tecnologia, anche l'approccio del medico è cambiato [65]. Oggi è possibile confermare o rivisitare diagnosi effettuate con metodi tradizionali e quindi predisporre per terapie più efficaci e mirate [42].

I gemelli, per le loro caratteristiche particolari, sono stati sempre oggetto di studio al fine di determinare il ruolo di fattori genetici ed ambientali nell'insorgenza e sviluppo di molteplici patologie [1, 6]. L'applicazione della tecnologia dei microarray allo studio dei gemelli è solo agli inizi, e nonostante ciò apprezzabili risultati sono stati raggiunti. L'analisi dell'espressione genica in coppie di gemelli MZ discordanti per patologie oncologiche e psichiatriche suggerisce che influenze ambientali sia nel periodo prenatale sia post-natale possano determinare differenze fenotipiche [57, 58]. La ricerca a riguardo si sta indirizzando verso lo studio di quelle modificazioni a carico del DNA definite epigenetiche. Queste interessano principalmente l'acetilazione degli istoni e la metilazione delle citosine e comportano un rimodellamento della cromatina e modificazioni conformazionali [56]. In ultima analisi questi ultimi eventi possono indurre alterazioni nei siti di legame per i fattori di trascrizione con conseguente silenziamento o attivazione dell'espressione genica [56, 63]. L'evoluzione della tecnologia dei microarray permette oggi l'analisi

di tutte queste modificazioni. In un prossimo futuro, quest'approccio su un campione sufficientemente ampio di gemelli renderà possibile definire il ruolo di queste variazioni del DNA nelle discordanze fenotipiche [60-62]. L'utilizzo dei gemelli MZ è determinante in questo tipo di analisi poiché le differenze osservate sono rilevate sullo sfondo di uno stesso substrato genetico [1-6]. La possibilità di disporre di gemelli MZ teoricamente identici sotto il profilo genetico, ma discordanti per patologie, rappresenta il sistema ideale per la comprensione di quei minimi cambiamenti che incidono sullo sviluppo di un fenotipo differente. Si può postulare che il connubio tra la tecnologia dei microarray ed il modello sperimentale dei gemelli potrà fornire, in un prossimo futuro, un importante contributo alla caratterizzazione dei processi biologici a livello molecolare e determinare le conoscenze basilari per una nuova medicina [10, 28]. I RG avranno un peso determinante, perché forniranno i dati sulle variabili fenotipiche e gli stili di vita da associare alle variazioni rilevate a livello molecolare [1-3]. L'insieme dei dati epidemiologici e molecolari potrà così aiutare a stabilire il ruolo dei fattori genetici ed ambientali nello sviluppo dei caratteri complessi, patologici e non.

Le informazioni acquisite tramite i Progetti Genoma e le nuove tecnologie pongono importanti interrogativi riguardo al loro uso e la loro divulgazione [66, 67]. A tale proposito, affinché le nuove conoscenze siano di beneficio per tutta l'umanità e le nuove tecnologie siano ispirate al rispetto della dignità dell'essere umano, è stato

sviluppato all'interno del PGU un programma di ricerca, denominato Ethical, Legal and Social Implications, che coinvolge, oltre a biologi molecolari e medici, anche avvocati, teologi, psicanalisti, sociologi ed altre figure professionali esperte di problematiche sociali e scienze umane. Pur riconoscendo l'importanza sociale di un corretto utilizzo delle risorse scientifiche derivate dal PGU, ci limitiamo a rimandare per questo argomento a testi specifici [66, 67].

In conclusione, i risultati già ottenuti dall'applicazione delle ricerche Genomica e Post-Genomica allo studio dei gemelli aprono nuove frontiere nell'epidemiologia genetica. Con l'utilizzo della tecnica dei microarray per lo studio dei gemelli si potranno acquisire preziose informazioni sui meccanismi genetico-molecolari di innumerevoli patologie umane. Queste informazioni potranno indirizzare la ricerca verso la messa a punto di nuovi strumenti di diagnosi e terapia.

### Ringraziamenti

Simonetta Pulciani e Anna Di Lonardo sono autori a pari partecipazione.

Si ringrazia il progetto GenomEUtwin (European Union Contract No. QL62-CT-2002-01254) per il supporto economico al lavoro di Corrado Fagnani e Cristina D'Ippolito. Un ringraziamento è rivolto anche ad Alessandro Spurio per l'assistenza tecnica nella realizzazione delle immagini.

Ricevuto il 27 ottobre 2005.

Accepted on 24 gennaio 2006.

### Bibliografia

- Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nat Rev Genet* 2002;3(11):872-82.
- Busjahn, A. (Ed.). Twin registers as a global resource for genetic research. *Twin Res* 2002;5(5):317-506.
- Peltonen L. GenomEUtwin: a strategy to identify genetic influences on health and disease. *Twin Res* 2003;6(5):354-60.
- Skytthe A, Pedersen NL, Kaprio J, Stazi MA, Hjelmberg JV, Iachine I, Vaupel JW, Christensen K. Longevity studies in GenomEUtwin. *Twin Res* 2003;6(5):448-54.
- Silventoinen K, Sarmalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, Dunkel L, De Lange M, Harris JR, Hjelmberg JV, Luciano M, Martin NG, Mortensen J, Nistico L, Pedersen NL, Skytthe A, Spector TD, Stazi MA, Willemssen G, Kaprio J. Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Res* 2003;6(5):399-408.
- Greco L, Romino R, Coto I, Di Cosmo N, Percopo S, Maglio M, Paparo F, Gasperi V, Limongelli MG, Cotichini R, D'Agate C, Tinto N, Sacchetti L, Tosi R, Stazi MA. The first large population based twin study of coeliac disease. *Gut* 2002;50(5):624-28.
- Khoury MJ, Beaty TH and Cohen BH. Fundamentals of genetic epidemiology. New York: Oxford University Press; 1993.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferreira S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooshep S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman

- M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. The sequence of the human genome. *Science* 2001;291(5507):1304-51.
9. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001; 409(6822):860-21.
  10. Khoury MJ, Davis R, Gwinn M, Lindegren ML, Yoon P. Do we need genomic research for the prevention of common diseases with environmental causes? *Am J Epidemiol* 2005;161(9):799-805.
  11. Cheung VG, Morley M, Aguilar F, Massimi A, Kucherlapati R, Childs G. Making and reading microarrays. *Nat Genet* 1999;21(1 Suppl):15-9.
  12. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 1975;98(3): 503-17.
  13. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;270(5235): 467-70.
  14. Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor, SPA. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* 1994;91(11):5022-6.
  15. McGall G, Labadie J, Brock P, Wallraff G, Nguyen T, Hinsberg W. Light-directed synthesis of high-density oligonucleotide arrays using semiconductor photoresists. *Proc Natl Acad Sci USA* 1996;93(24):13555-60.
  16. Okamoto T, Suzuki T, Yamamoto N. Microarray fabrication with covalent attachment of DNA using bubble jet technology. *Nat Biotechnol* 2000;18(4):438-41.
  17. Auburn RP, Kreil DP, Meadows LA, Fischer B, Matilla SS, Russell S. Robotic spotting of cDNA and oligonucleotide microarrays. *Trends Biotechnol* 2005;23(7):374-9.
  18. Gaasterland T, Bekiranov S. Making the most of microarray data. *Nat Genet* 2000;24(3):204-6.
  19. Cavalieri D, De Filippo C. Bioinformatic methods for integrating whole-genome expression results into cellular networks. *Drug Discov Today* 2005;10(10):727-34.
  20. Pinkel D, Albertson DG. Array comparative genomic hybridization and its applications in cancer. *Nat Genet* 2005;37(Suppl):S11-7.
  21. Yamamoto F, Yamamoto M. A DNA microarray-based methylation-sensitive (MS)-AFLP hybridization method for genetic and epigenetic analyses. *Mol Genet Genomics* 2004;271(6):678-86.
  22. van Steensel B. Mapping of genetic and epigenetic regulatory networks using microarrays. *Nat Genet* 2005;37(Suppl):S18-24.
  23. Watanabe A, Cornelison R, Hostetter G. Tissue microarrays: applications in genomic research. *Expert Rev Mol Diagn* 2005;5(2): 171-81.
  24. Eckel-Passow JE, Hoering A, Therneau TM, Ghobrial I. Experimental design and analysis of antibody microarrays: applying methods from cDNA arrays. *Cancer Res* 2005;65(8):2985-9.
  25. Quackenbush J. Extracting meaning from functional genomics experiments. *Toxicol Appl Pharmacol* 2005;207(2 Suppl):195-9.
  26. Li X, Quigg RJ. An integrated strategy for the optimization of microarray data interpretation. *Gene Expr* 2005;12(4-6):223-30.
  27. Stoughton RB. Applications of DNA microarrays in biology. *Annu Rev Biochem* 2005;74:53-82.
  28. Collins FS, Green ED, Guttmacher AE, Guyer MS. US National Human Genome Research Institute. A vision for the future of genomics research. *Nature* 2003;422(6934):835-47.
  29. Iyer VR. Exploring the post-transcriptional RNA world with DNA microarrays. *Trends Biotechnol* 2004;22(10):498-500.
  30. Pollack JR, Iyer VR. Characterizing the physical genome. *Nat Genet* 2002;32 (Suppl):515-21.
  31. Hacia JG, Collins FS. Mutational analysis using oligonucleotide microarrays. *J Med Genet* 1999;36(10):730-6.
  32. Ge H, Walhout AJ, Vidal M. Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Genet* 2003;19(10):551-60.
  33. Schena M, Shalon D, Heller R, Chai A, Brown PO, and Davis RO. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci USA* 1996;93 (20):10614-9.
  34. Sacha T, Zawada M, Hartwich J, Lach Z, Polus A, Szostek M, Zdzi Owska E, Libura M, Bodzioch M, Dembinska-Kiec A, Skotnicki AB, Goralczyk R, Wertz K, Riss G, Moele C, Langmann T, Schmitz G. The effect of beta-carotene and its derivatives on cytotoxicity, differentiation, proliferative potential and apoptosis on the three human acute leukemia cell lines: U-937, HL-60 and TF-1. *Biochim Biophys Acta* 2005;1740(2):206-14.
  35. Giordano TJ, Kuick R, Thomas DG, Misesk DE, Vinco M, Sanders D, Zhu Z, Ciampi R, Roh M, Shedden K, Gauger P, Doherty G, Thompson NW, Hanash S, Koenig RJ, Nikiforov YE. Molecular classification of papillary thyroid carcinoma: distinct BRAF, RAS, and RET/PTC mutation-specific gene expression profiles discovered by DNA microarray analysis. *Oncogene* 2005;24(44):6646-56.
  36. Calvo A, Gonzalez-Moreno O, Yoon CY, Huh JI, Desai K, Nguyen QT, Green JE. Prostate cancer and the genomic revolution: Advances using microarray analyses. *Mutat Res* 2005;576(1-2): 66-79.
  37. Bryant PA, Venter D, Robins-Browne R, Curtis N. Chips with everything: DNA microarrays in infectious diseases. *Lancet Infect Dis* 2004;4(2):100-11.
  38. Bacarese-Hamilton T, Mezzasoma L, Ardizzone A, Bistoni F, Crisanti A. Serodiagnosis of infectious diseases with antigen microarrays. *J Appl Microbiol* 2004;96(1):10-7.
  39. Nadon NL, Mohr D, Becker KG. National Institute on Aging microarray facility- resources for gerontology research. *J Gerontol A Biol Sci Med Sci* 2005;60(4):413-5.
  40. Kallioniemi OP. Biochip technologies in cancer research. *Ann Med* 2001;33(2):142-7.
  41. Kleivi K, Diep CB, Pandis N, Heim S, Teixeira MR, Lothe RA. TP53 mutations are associated with a particular pattern of genomic imbalances in breast carcinomas. *J Pathol* 2005;207(1):14-9.
  42. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286(5439):531-7.
  43. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JJ, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403(6769):503-11.
  44. Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, Tibshirani R, Botstein D, Borresen-Dale AL, Brown PO. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl. Acad. Sci USA* 2002;99(20):12963-8.
  45. Chan VL. Related Bacterial genomes and infectious diseases. *Pediatr Res* 2003;54(1):1-7.
  46. Robertson BH, Nicholson JK. New microbiology tools for public health and their implications. *Annu Rev Public Health* 2005;26: 281-302.
  47. Weindruch R, Kayo T, Lee CK, Prolla TA. Microarray profiling of gene expression in aging and its alteration by caloric restriction in mice. *J Nutr* 2001;131(3):918S-23S.

48. Lee CK, Klopp RG, Weindruch R, Prolla TA. Gene expression profile of aging and its retardation by caloric restriction. *Science* 1999;285(5432):1390-3.
49. Devlin B, Roeder K, Bacanu SA. Unbiased methods for population-based association studies. *Genet Epidemiol* 2001;21(4):273-84.
50. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* 2005;33:D514-7.
51. Mayeux R. Mapping the new frontier: complex genetic disorders. *J Clin Invest* 2005;115(6):1404-7.
52. Dalma-Weiszhausz DD, Chicurel ME, Gingeras TR. Microarrays and genetic epidemiology: a multipurpose tool for a multifaceted field. *Genet Epidemiol* 2002;23(1):4-20.
53. Munshi NC, Hideshima T, Carrasco D, Shamma M, Auclair D, Davies F, Mitsiades N, Mitsiades C, Kim RS, Li C, Rajkumar SV, Fonseca R, Bergsagel L, Chauhan D, Anderson KC Identification of genes modulated in multiple myeloma using genetically identical twin samples. *Blood* 2004;103(5):1799-806.
54. Kakiuchi C, Iwamoto K, Ishiwata M, Bundo M, Kasahara T, Kusumi I, Tsujita T, Okazaki Y, Nanko S, Kunugi H, Sasaki T, Kato T. Impaired feedback regulation of XBP1 as a genetic risk factor for bipolar disorder. *Nat Genet* 2003;35(2):171-5.
55. Kelsoe JR, Spence MA, Loetscher E, Foguet M, Sadovnick AD, Remick RA, Flodman P, Khristich J, Mroczkowski-Parker Z, Brown JL, Masser D, Ungerleider S, Rapaport MH, Wishart WL, Luebbert H. A genome survey indicates a possible susceptibility locus for bipolar disorder on chromosome 22. *Proc Natl Acad Sci USA* 2001;98(2):585-90.
56. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 2004;429(6990):457-63.
57. Petronis A, Gottesman II, Kan P, Kennedy JL, Basile VS, Paterson AD, Popenklyte V. Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance? *Schizophr Bull* 2003;29(1):169-78.
58. Dennis C. Epigenetics and disease: Altered states. *Nature* 2003;421(6924):686-8.
59. Yasui W, Oue N, Aung PP, Matsumura S, Shutoh M, Nakayama H. Molecular-pathological prognostic factors of gastric cancer: a review. *Gastric Cancer* 2005;8(2):86-94.
60. Tan Q, Christensen K, Christiansen L, Frederiksen H, Bathum L, Dahlgaard J, Kruse TA. Genetic dissection of gene expression observed in whole blood samples of elderly Danish twins. *Hum Genet* 2005;117(2-3):267-74.
61. Sharma A, Sharma VK, Horn-Saban S, Lancet D, Ramachandran S, Brahmachari SK Assessing natural variations in gene expression in humans by comparing with monozygotic twins using microarrays. *Physiol Genomic* 2005;21(1):117-23.
62. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, Heine-Suner D, Cigudosa JC, Urioste M, Benitez J, Boix-Chornet M, Sanchez-Aguilera A, Ling C, Carlsson E, Poulsen P, Vaag A, Stephan Z, Spector TD, Wu YZ, Plass C, Esteller M. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci USA* 2005;102(30):10604-9.
63. York TP, Miles MF, Kendler KS, Jackson-Cook C, Bowman MI and Eaves LJ Epistatic and environmental control of genome-wide gene expression. *Twin Res Hum Genet* 2005;8:5-15.
64. Angenendt P. Progress in protein and antibody microarray technology. *Drug Discov Today* 2005;10(7):503-11.
65. Weeraratna AT, Nagel JE, de Mello-Coelho V, Taub DD. Gene expression profiling: from microarrays to medicine. *J Clin Immunol* 2004;24(3):213-24.
66. Harris JR, Willemsen G, Aitlahti T, Petrini C, Evans A, Silander K, Cirrincione L, Kyvik KO. Ethical issues and GenomEUtwin. *Twin Res* 2003;6(5):455-63.
67. Cho MK, Sankar P. Forensic genetics and ethical, legal and social implications beyond the clinic. *Nat Genet* 2004;36(Suppl):8-12.