

## INTERAZIONE GENOMA AMBIENTE

Simona Gaudi

Dipartimento Ambiente e Salute, Istituto Superiore di Sanità, Roma

### Nuova concezione del genoma

Il genoma è l'insieme delle informazioni ereditarie di un organismo; nel genoma risiedono sia le sequenze che codificano per le proteine, i “mattoni” per la costruzione di un organismo, che sequenze apparentemente non codificanti per proteine, erroneamente ritenute senza una funzione (Lander *et al.*, 2001; Venter *et al.*, 2001). In realtà queste sequenze, che costituiscono più del 98% dell'intero genoma umano, sono alla base della regolazione dell'espressione delle proteine. Le tecnologie disponibili, prima dello sviluppo del Progetto Genoma Umano, consentivano l'identificazione e l'analisi solo di alcuni geni ipoteticamente coinvolti nell'eziopatogenesi delle malattie. È necessario, oggi rivedere il concetto di DNA spazzatura o egoista, e riassegnare una funzione a tali sequenze che costituiscono la base della nostra unicità e che potrebbero essere coinvolte nella predisposizione alle malattie multigeniche e multifattoriali. Uno dei risultati più sorprendenti e inattesi ottenuti dal progetto di sequenziamento del genoma umano è stato proprio quello di constatare che il numero totale dei geni è decisamente inferiore a quello atteso e sorprendentemente simile a quello di organismi molto diversi: un semplice nematode come il *C. elegans* ha un numero di geni codificanti (~19,300) quasi uguale a quello umano (~20,000) (Di Leonardo *et al.*, 2009). Questo ha suggerito che il grado di complessità biologica di un organismo non possa essere semplicemente il risultato del numero dei geni presenti nel suo DNA, ma piuttosto, dipenda dal grado di complessità dei meccanismi che ne regolano l'espressione (Copley, 2008; Stumpf *et al.*, 2008). Non dobbiamo dimenticare che tutte le componenti del genoma sono fondamentali e solo un corretto funzionamento e regolazione può garantirci lo stato di salute. La conoscenza dell'esatta sequenza dei nostri geni non è sufficiente a spiegare il confine che esiste tra stato di salute e malattia. Quindi, il nostro genoma nella sua complessità complessa, è dotato di una plasticità che gli consente di interagire con l'ambiente e di modificarsi nel tempo.

Il sequenziamento completo del genoma umano e le tecnologie di analisi funzionale del genoma stesso, hanno consentito lo studio delle variazioni del DNA grazie all'esistenza di polimorfismi genetici (*Single Nucleotide Polymorphism*, SNP), dell'intero “trascrittoma” (tutto ciò che l'intero genoma trascrive, RNA) e del proteoma (tutte le proteine tradotte) al fine di individuare tutte le componenti molecolari che dipendono dal genoma.

### Patologie non trasmissibili

Tali progressi della ricerca scientifica e le conoscenze di genomica funzionale hanno permesso di studiare e parzialmente decodificare la natura complessa di molte malattie non trasmissibili. Le patologie non trasmissibili ovvero multigeniche, multifattoriali e croniche, sono il risultato dell'interazione tra fattori biologici, psicologici, sociali e ambientali. Sebbene tutte le malattie non trasmissibili, definite anche come croniche e multifattoriali, hanno una componente genetica, l'identificazione di tutti i geni responsabili rappresenta ancora oggi una delle sfide più granci della medicina di precisione. Come mai siamo ancora molto lontani dal potere individuare i geni

responsabili? Forse questo può dipendere dal fatto che la ricerca genetica degli ultimi vent'anni si è concentrata essenzialmente, se non esclusivamente, su quella porzione del genoma (2%) che codifica per le proteine. Lo studio dell'interazione non solo dei geni, ma dell'intero genoma con i diversi fattori ambientali rappresenta invece oggi una delle aree più promettenti per la medicina di precisione. Studi di genomica funzionale e di sequenziamento completo del genoma, tecnologie impensabili fino a qualche anno fa, offrono e offriranno nuove prospettive per comprendere le basi molecolari dell'interazione genoma/ambiente nelle patologie complesse.

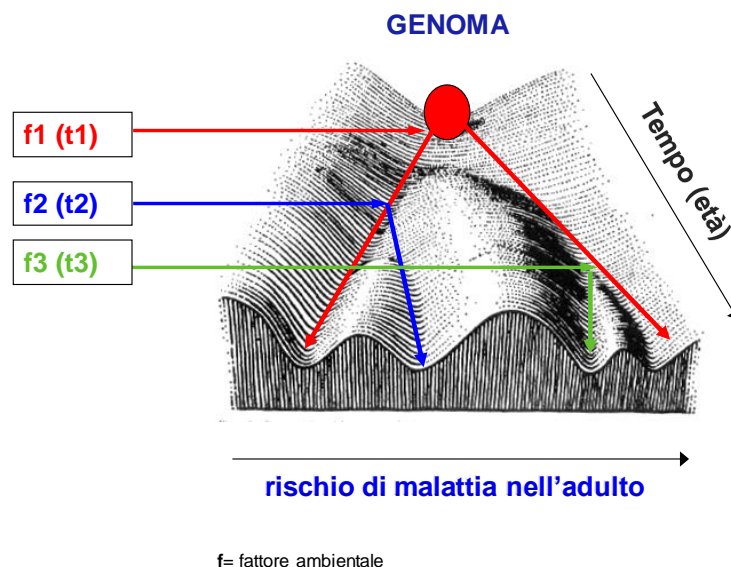
## Epigenetica

La letteratura scientifica degli ultimi 20-30 anni ha dimostrato come fattori ambientali siano in grado di alterare la struttura e la funzionalità del nostro genoma. Gli effetti dell'ambiente sul genoma possono influenzare la struttura del DNA compromettendo lo stato di salute dell'individuo. Le modificazioni epigenetiche producono un rimodellamento della struttura della cromatina e modulano l'espressione dei geni adiacenti. Tali meccanismi vengono definiti epigenetici in quanto, pur non producendo alterazioni stabili nel codice del DNA, sono ugualmente in grado di indurre modificazioni trasmissibili del fenotipo (Weinhold, 2006). Le modificazioni e le variazioni epigenetiche sono un meccanismo naturale che consente allo stesso genoma di essere letto in modo differente dando origine ai diversi tessuti e di conseguenza ai differenti organi. Se non intervenisse la regolazione epigenetica, la formazione e lo sviluppo di un embrione in un organismo completo non potrebbe avvenire. Ad oggi si conoscono tre meccanismi in grado di regolare la lettura delle informazioni del filamento di DNA, metilazione delle basi, acetilazione degli istoni e la presenza di piccole molecole di RNA che stabilizzano la struttura del DNA. Lo studio dei marcatori epigenetici (modificazioni molecolari e chimiche che non comportano cambiamenti nel nostro codice genetico), potrebbe aiutarci a capire quali sono le cause delle malattie complesse e, quindi, pensare a una loro possibile modificazione per attenuare o rimuovere le conseguenze biologiche in seguito all'interazione con fattori ambientali negativi.

Il meccanismo più studiato è la metilazione, ovvero l'aggiunta di un gruppo metile CH<sub>3</sub>, a livello delle regioni ricche in citosina che sono definite come isole di CpG (*CpG-islands*). La metilazione reprime la trascrizione sia direttamente, inibendo il legame dei fattori di trascrizione, sia indirettamente, reclutando proteine che legano le sequenze CpG-metilate. Altri meccanismi di controllo epigenetico comportano modificazioni covalenti della coda degli istoni, come l'acetilazione, la fosforilazione, l'ubiquitinazione. Queste modificazioni alterano la struttura degli istoni, influenzando l'organizzazione dei complessi cromatinici e quindi l'accessibilità all'informazione genetica (Gopalakrishnan *et al.* 2008, Hadnagy *et al.*, 2008). In generale, si può dire che quanto più la cromatina è condensata tanto più la trascrizione risulta repressa. Un altro meccanismo epigenetico è costituito da alcune forme di RNA, come i miRNA (micro RNA) e i siRNA (*small interfering RNA*). Si tratta di piccole molecole di 21-25 nucleotidi che agiscono come "reostati cellulari" (modulatori) nella regolazione fine dell'espressione genica durante lo sviluppo e la differenziazione che hanno come bersaglio il 3' UTR (UnTranslated Region) degli mRNA, regioni coinvolte nell'espressione, stabilizzazione, localizzazione ed efficienza traduzionale degli RNA messaggeri (Wang *et al.*, 2008; Kim & Rossi, 2008). I meccanismi epigenetici regolano dunque i "pattern" di espressione che sono alla base dello sviluppo embrionale, del differenziamento cellulare, della risposta delle cellule a segnali ambientali come ormoni, nutrienti, stress, invecchiamento e danno cellulare. L'innovazione tecnologica di questi ultimi anni ha trasformato e ampliato le conoscenze nel campo dell'epigenetica. Tecniche quali i *microarray*, *ultra-high-throughput sequencing*, e altre ancora, sono state utilizzate per mappare nel genoma le modificazioni della cromatina, la metilazione delle cisteine e gli RNA non

codificanti, con una risoluzione al livello di proteosoma, dell'ordine di poche centinaia di basi. Questa alta capacità di risoluzione sta già permettendo di identificare differenze tra l'epigenoma di tessuti primari normali e campioni patologici, così come fra le diverse popolazioni cellulari che compongono un organismo. L'epigenoma integra le informazioni contenute nel genoma con tutti i segnali molecolari e chimici di origine cellulare, extracellulare e ambientale. Insieme al genoma, l'epigenoma impartisce il programma di espressione genico specifico e unico di ciascun tipo di cellula per definire la sua identità funzionale durante lo sviluppo o la malattia (Rivera *et al.*, 2013). L'epigenoma rappresenta anche la capacità di un organismo di adattarsi ed evolvere attraverso l'espressione di un set di caratteristiche o fenotipi che si sono sviluppati in risposta ad uno stimolo ambientale. In contrasto con la minore plasticità delle sequenze del genoma, l'epigenoma è caratterizzato da una risposta dinamica e flessibile a stimoli intracellulari ed extracellulari attraverso contatto tra cellule vicine, a cambiamenti fisiologici o a cambiamenti ambientali a cui l'organismo è esposto.

Tutti i geni in tutti i tipi di cellule sono attivati o silenziati attraverso l'interazione di questi meccanismi epigenetici. Durante la vita di un individuo sono presenti moltissimi cambiamenti epigenetici che si integrano cronologicamente con i diversi fattori ambientali che interagiscono con l'epigenoma. Nelle cellule differenziate sono presenti segnali che regolano in modo fine le funzioni cellulari durante la vita. Un epigenoma flessibile permette all'individuo di modificarsi per far fronte ai cambiamenti dell'ambiente circostante e imparare dalle nostre esperienze. L'espressione epigenetica può essere considerata come "il software" (Kanherkar *et al.*, 2014) del genoma che dirige l'embriogenesi e lo sviluppo prenatale e dirige anche lo sviluppo del corpo e del cervello di un individuo dopo la nascita. Specifici gruppi di geni vengono attivati o silenziati epigeneticamente durante diversi periodi della vita e questi sono responsabili dello sviluppo e della maturazione dell'individuo attraverso eventi predisposti in combinazione con stimoli ambientali. Ogni tipo di fattore epigenetico che influenza i geni o il network di espressione dei geni durante gli stadi della vita può provocare uno squilibrio nel processo di regolazione e può avere anche effetti a lungo termine (Figura 1) (Taherian & Ragan, 2019)



**Figura 1. Rappresentazione di come da un'unico genoma, grazie anche all'interazione con fattori ambientali, si ottengono epigenomi diversi**

## Conclusioni

I più recenti risultati dimostrano come tutto il genoma concorra nel determinare le tante e diverse malattie complesse indicando un nuovo percorso per la ricerca delle basi molecolari delle malattie non trasmissibili. Lo studio della genomica funzionale e del sequenziamento completo del genoma ci consente di interpretare la malattia come la risultante di una variazione nella regolazione di più fattori: geni, genoma, epigenoma e ambiente.

La vulnerabilità genetica e genomica porta, infatti, allo stato di malattia quando si scontra con aspetti incongruenti dell'ambiente, ovvero delle abitudini e degli stili di vita dell'individuo. Allora è chiaro che, attraverso la conoscenza del genoma umano si identificheranno le basi della vulnerabilità verso le malattie, e di conseguenza l'intervento sull'ambiente e sugli stili di vita potrà diventare sempre più determinante per la cura e la prevenzione delle malattie non trasmissibili. Approcci innovativi per lo studio di tali meccanismi potranno aiutarci nell'individuare marker molecolari precoci, anche a livello del sangue periferico, che siano informativi e predittivi dell'insorgenza di una patologia.

L'elaborazione delle conoscenze scientifiche, che interpretano con sempre maggiore accuratezza, la reale complessità biologica delle malattie non trasmissibili, dovrà tradursi in pratiche cliniche in grado rispondere tempestivamente all'esordio delle patologie.

In questo contesto, la ricerca deve assumere una dimensione più integrata, fondata sul continuo interscambio di esperienze, tecniche e dati da ottenere con il potenziamento delle relazioni che verranno acquisite. Solo lo studio dell'architettura e della variazione del genoma umano, in relazione al cambiamento dei fattori ambientali, consentirà di comprendere non solo la storia genetica delle popolazioni, ma soprattutto le radici di molte patologie.

## Bibliografia

- Copley R. The animal in the genome: comparative genomics and evolution. *Philos Trans R Soc Lond B Biol Sci* 2008;363(1496):1453-61.
- Di Leonardo A, Fagnani C, Pulciani S (Ed.). Applicazione dei microarray (2). *Caleidoscopio italiano* 2009;223.
- Gopalakrishnan S, Van Emburgh BO, Robertson KD. DNA methylation in development and human disease. *Mutat Res* 2008;647(1-2):30-8.
- Hadnagy A, Beaulieu R, Balicki D. Histone tail modifications and non canonical functions of histones: perspectives in cancer epigenetics. *Mol Cancer Ther* 2008;7(4):40-8.
- Kanherkar RR, Bhatia-Dey N, Csoka AB. Epigenetics across the human lifespan. *Front Cell Dev Biol* 2014;2:49.
- Kim D, Rossi J. RNAi mechanisms and applications. *Biotechniques* 2008;44(5): 613-6.
- Lander E, *et al.* International human genome sequencing consortium. Finishing the eucromatic sequence of the human Genome. *Nature* 2004;431:931-45.
- Rivera CM, Ren B. Mapping human epigenomes. *Cell* 2013; 155(1): 39-55.
- Stumpf MP., Thorne T, de Silva, Stewart R, An HJ, Lappe M, Wiuf C. Estimating the size of the human interactome. *Proc Natl Acad Sci USA* 2008; 105(19): 6959-64.
- Taherian Fard A, Ragan MA. Quantitative modelling of Waddington epigenetic landscape. *Methods Mol Biol* 2019;1975:157-171.
- Venter JC, Adams MD, Meyers EW, Li PW, Mural RJ, Sutton GG *et al.* The sequence of the human genome. *Science* 2001; 291(5507): 1304-51.

Wang Y, Liang Y, Lu Q. MicroRNA epigenetic alterations: predicting biomarkers and therapeutic targets in human diseases. *Clin Genet* 2008;74(4):307-15.

Weinhold B. Epigenetics: the science of change. *Environ Health Perspect* 2006;114(3):A 160-7.