EPIGENETICA (2014)

Il "dogma centrale" della biologia molecolare sostiene che le informazioni ereditarie sono trasmesse attraverso meccanismi genetici.

In realtà, lungo le generazioni, una cellula scambia con le cellule figlie anche informazioni non contenute nella sequenza di basi del DNA. L'**epigenetica** studia la **trasmissione di caratteri ereditari non attribuibili direttamente alla sequenza di DNA.** Un paragone esemplificativo potrebbe essere quello che considera il cromosoma come un libro con un testo completo e corretto (la sequenza di basi), ma fattori esterni al "testo", come l'incollamento di alcune pagine tra loro, potrebbero non permettere l'accesso alle informazioni contenute nel testo stesso.

1. Metilazione del DNA

Un evento molecolare noto che ostacola la "lettura" delle basi, ossia l'espressione genica, è la **metilazione del DNA.** Gli enzimi della classe delle **DNA metil-transferasi** trasferiscono gruppi metilici al DNA. Si pensa che la metilazione di un gene, in particolare nella sua regione regolativa, ne influenzi l'espressione. Ad esempio è stata dimostrata una proporzionalità diretta tra la metilazione di un promotore e la mancata espressione del gene a valle. Lo **schema della metilazione del DNA** viene trasmesso lungo le generazioni cellulari successive. Se si considera un individuo omozigote per un gene ma in cui i due alleli differiscono per lo schema di metilazione si avrà che uno dei due alleli potrà essere più espresso dell'altro. La conseguenza sarà che, anche partendo da due regioni di DNA di sequenza identica si troverà una differenza funzionale, come se si fosse in presenza di due alleli diversi. Dato che l'allele viene trasmesso ereditariamente e dato che nel passaggio da una cellula ad un'altra un gene tende a mantenere il proprio schema di metilazione, si ottiene complessivamente la trasmissione di una informazione sull'espressione del gene associata al gene stesso ma indipendente dalla sua sequenza nucleotidica.

2. Organizzazione della cromatina

Un altro aspetto che influenza la trascrizione è la **configurazione della cromatina.** Il cromosoma può assumere diversi gradi di compattamento, e i gradi superiori di addensamento ostacolano la trascrizione del gene. Quindi per studiare il funzionamento del gene è necessario conoscere la configurazione locale della cromatina. Anche la configurazione della cromatina non dipende direttamente dalla sequenza del DNA, ma da come i meccanismi cellulari in una certa fase funzionale hanno determinato il ripiegamento dei cromosomi nucleari.

3. Imprinting genomico

Un terzo meccanismo di regolazione della trascrizione di origine epigenetica è l'**imprinting genomico.** Con imprinting si intende il fenomeno per cui uno dei due alleli di un gene è escluso selettivamente dall'attivazione, in dipendenza dal genitore (madre o padre) da cui il cromosoma recante l'allele stesso è stato ricevuto (**esclusione allelica dipendente dalla origine parentale**).

Consideriamo, ad esempio, il gene della globina, sul cromosoma 11; si assume che il gene materno e quello paterno produrranno la stessa quantità di mRNA, che sarà tradotto in proteina, e quindi circa metà della proteina circolante deriverà dall'allele materno il restante dall'allele paterno. In alcune

circostanze particolari solo un allele funziona, mentre la trascrizione dell'altro risulta inibita (**esclusione allelica**).

Esempio di geni con meccanismi di esclusione fisiologicamente attivi sono quelli per gli **anticorpi**, nel caso dei quali è attivo o il locus materno o quello paterno, oppure i geni della copia del cromosoma **X** che viene inattivata nelle cellule femminili.

Altri esempi di espressione selettiva di un allele vengono dalla patologia, nel caso in cui, in presenza di una mutazione di uno dei due alleli che ne impedisca l'espressione (ad esempio per lesione della regione regolativa), viene espresso solo l'allele normale.

In un numero piccolo, ma non trascurabile, di loci umani (oggi stimato intorno a 100), è stato descritto recentemente un importante nuovo meccanismo di esclusione allelica. Questi geni sono espressi selettivamente in tutti gli individui sempre a partire da uno dei due alleli, che per alcuni loci è sempre quello localizzato sul cromosoma derivato dal padre, per altri loci è sempre quello che si trova sul cromosoma di origine materna. Deve esistere quindi un fattore che permette alla cellula di distinguere l'allele materno da quello paterno e attivare solo o l'uno o l'altro, in modo specifico per quel dato gene. Per quanto detto a proposito della esclusione allelica dovuta a mutazioni, occorre precisare che per stabilire che un locus è soggetto fisiologicamente a imprinting i due alleli devono essere di sequenza normale o meglio ancora identica.

Un gene è sottoposto a **imprinting materno** quando viene inibita selettivamente l'espressione dell'allele materno e la proteina viene prodotta solo a partire dall'allele paterno. Viceversa, esiste il caso in cui viene represso l'allele di origine paterna (**imprinting paterno**) e la proteina viene sintetizzata solamente a partire dallo stampo costituito dall'allele materno. Occorre chiarire che **questo meccanismo di regolazione è attuato solo per un certo locus**, il gene adiacente ad uno inattivato per imprinting può essere regolato in modo del tutto normale o sottoposto ad un imprinting di segno opposto.

Nei **geni** sottoposti a imprinting, quindi sarà sempre espresso o l'allele materno o paterno. Nel caso in cui siano presenti, in un locus regolato attraverso l'imprinting, un allele normale soggetto ad imprinting e un allele mutato, ad esempio per una delezione, non soggetto ad imprinting, si osserverà nell'individuo che possiede questi alleli un fenotipo patologico. Viceversa, una mutazione in un gene che non sarà espresso perché sottoposto a imprinting potrebbe passare inosservata.

Stabilire che tutto l'RNA cellulare di un gene deriva da un solo allele può essere molto difficile, e di solito fattibile solo in modelli animali, se i due alleli hanno sequenza identica; il compito è più facile se i due alleli si differenziano almeno per una coppia di basi.

Esempi di geni umani repressi per imprinting sono: l'**Insulin-like growth factor 2**, *IGF2*, un gene localizzato sul cromosoma 11 di cui viene represso selettivamente l'allele materno, e il cui mRNA deriva quindi dall'espressione del solo allele paterno, o il gene adiacente chiamato ***H19***, codificante per un RNA non tradotto e sottoposto viceversa a imprinting paterno, e di cui viene quindi espresso solo l'allele di origine materna.

Il fenomeno dell'imprinting non è stato descritto in tutti gli organismi. Attualmente è stato dimostrato in semi di piante, in alcuni insetti e nei mammiferi, mentre non si ritrova in altri invertebrati o vertebrati: il moscerino della frutta *Drosophila melanogaster*, il verme nematode *C. elegans*, il pesce zebra (*Danio rerio*). Tale fenomeno non è, di conseguenza, indispensabile per la vita, ma nelle specie in cui si verifica è indispensabile per il corretto funzionamento del genoma.

Il **meccanismo molecolare** con cui la cellula realizza il fenomeno dell'imprinting è tuttora sconosciuto, sebbene sia correlato probabilmente alla **metilazione** del DNA. La repressione di un allele potrebbe avvenire tramite un alto grado di metilazione selettiva.

Nell'uomo la maggioranza dei geni soggetti a imprinting è autosomica. Molti dei geni umani soggetti a imprinting si trovano raggruppati in due localizzazioni principali nel genoma: una zona di circa 1 Mbp si trova sul cromosoma 11 (11p15.5, ossia: braccio corto del cromosoma 11, regione 1, banda 5, sottobanda 5), mentre un secondo raggruppamento, esteso per 2,3 Mbp, si trova sul cromosoma 15 (15q11-q13).

Alcuni di questi geni sono coinvolti nella regolazione della crescita del feto: i geni paterni partecipano a rimuovere i nutrienti in modo aggressivo dal corpo della madre, per cui sono espressi nel trofoblasto e nelle membrane extraembrionarie; i geni materni si oppongono all'effetto dei geni paterni, "proteggono la madre" provocando l'accumulo di nutrienti per le cellule discendenti, e sono espressi soprattutto nell'embrione. Per alcuni geni l'imprinting è confinato solo in certi **tessuti** o in determinati **stadi dello sviluppo**. In questo senso l'imprinting permette un livello aggiuntivo di controllo dell'espressione genica, di cui però, è necessario chiarire molti meccanismi. Ad esempio, il gene citato *IGF2* è sottoposto a imprinting nella maggior parte dei tessuti, ma nel cervello e in pochi altri distretti mostra espressione biallelica.

TRASMISSIONE DEI GENI SOTTOPOSTI A IMPRINTING

La trasmissione da una generazione all'altra di geni con imprinting richiede che debba esistere un meccanismo di "riazzeramento" del meccanismo molecolare dell'imprinting e uno specifico momento in cui il gene viene "marcato" in modo sesso-specifico con l'imprinting. Un padre deve trasmettere un gene con imprinting comunque di tipo paterno, anche se lui stesso lo ha ricevuto dalla madre, quindi con imprinting materno.

Nella linea germinale umana è stata individuata una specifica **metilasi** che viene espressa, in quantità abbondanti, solo durante la maturazione degli ovociti, e corrisponderebbe, quindi, al profilo cercato per l'attuatore dell'imprinting sui geni materni. Nella linea germinale maschile probabilmente l'imprinting viene stabilito prima della meiosi, forse allo stadio di spermatocita primario postmitotico. L'azzeramento dell'imprinting potrebbe essere uno degli eventi che accadono precocemente, al momento della definizione delle linee germinali degli organismi.

Presentiamo uno schema che permette di **seguire l'imprinting dei geni *IGF2* e *H19***, sottoposti a imprinting di segno opposto, lungo due generazioni successive; l'allele represso, che ha subìto l'imprinting, è marcato da un asterisco:

```
Padre                    IGF2,  H19* (cromosoma paterno)
                         IGF2*, H19  (cromosoma materno)

Madre                    IGF2,  H19* (cromosoma paterno)
                         IGF2*, H19  (cromosoma materno)

Durante lo sviluppo dei gameti:

Spermio                  IGF2,  H19* (imprinting rimosso dall'allele IGF2*
                                      e stabilito sull'allele H19)

Ovocita                  IGF2*, H19  (imprinting rimosso dall'allele H19*
                                      e stabilito sull'allele IGF2)

Nuovo individuo          IGF2,  H19* (cromosoma paterno)
                         IGF2*, H19  (cromosoma materno)
```

```
Durante lo sviluppo dei gameti del nuovo individuo, se maschio:

Spermio                  IGF2,  H19* (imprinting rimosso dall'allele IGF2
                                      e stabilito sull'allele H19)

Se femmina:

Ovocita                  IGF2*, H19 (imprinting rimosso dall'allele H19*
                                      e stabilito sull'allele IGF2)

ecc.
```

Una dimostrazione dell'importanza dell'imprinting è dato da un articolo di recente pubblicazione. In questo lavoro si riporta il caso di una coppia di gemelli monovulari, uno dei quali risulta affetto da una malattia genetica e l'altro no. La spiegazione va cercata nel fatto che la metilazione dei geni inattivati dall'imprinting viene "rinforzata" durante le prime fasi della crescita degli zigoti. il mantenimento dell'imprinting corretto, dopo la scissione dello zigote originario nei due embrioni, può essere carente in uno dei due gemelli (si veda il capitolo sulla Eredità multifattoriale).

4. Conformazione della proteina

Un caso diverso della regolazione dell'espressione si può avere dopo la traduzione della proteina. Una proteina neosintetizzata assume una struttura terziaria in maniera dipendente dall'ambiente in cui si trova; se un certo ambiente, ad esempio la concentrazione salina di un compartimento cellulare, costringe questa proteina ad assumere una conformazione meno funzionale di un altro ambiente, si otterrà alla fine una situazione funzionalmente diversa a partire dalla stessa informazione iniziale, la sequenza amminoacidica della proteina determinata su base genetica. Anche l'informazione della conformazione finale che può assumere una proteina è trasmessa in maniera indipendente dalla sequenza di DNA che codifica per la proteina stessa.

Le proteine prioniche alterate, responsabili tra l'altro del morbo della mucca pazza e di alcune malattie umane, si replicano attraverso una "memoria" molecolare di tipo epigenetico che risiede appunto nella struttura tridimensionale alterata della proteina. Il termine **"prione"** è stato introdotto da Prusiner nel 1982 per indicare una nuova classe di agenti infettivi caratterizzata dalla mancanza di acido nucleico (DNA o RNA). La radice "pri" del nome è una abbreviazione dell'espressione "proteina infettiva", concetto che suonava all'epoca come un'eresia, in quanto si accettava che gli agenti infettivi che trasmettono le malattie dovessero comunque disporre di materiale genetico, in forma appunto di DNA o RNA, per essere in grado di replicarsi e quindi di propagare l'infezione in un ospite. Il prione fu identificato studiando un gruppo di malattie letali, denominate nel loro complesso encefalopatie spongiformi ("a forma di spugna"), in quanto queste malattie sono caratterizzate dalla formazione di bolle e cavità nella corteccia cerebrale, con conseguente danneggiamento progressivo delle funzioni nervose. La forma che colpisce pecore e capre è nota come scrapie, ed è caratterizzata da perdita della coordinazione muscolare, degenerazione progressiva del sistema nervoso centrale e prurito cronico e intenso, in conseguenza del quale gli animali colpiti si grattano contro oggetti fino a raschiar via (scrape, in inglese) parti del mantello. Prusiner trovò che l'agente infettivo dello scrapie era costituto prevalentemente, se non esclusivamente, da materiale proteico.
Il modello attualmente proposto per il funzionamento del prione è il seguente:
• la **proteina prionica** è una proteina naturalmente presente nell'organismo, e nell'uomo è costituita da una catena di 253 amminoacidi. Il **gene** che codifica per la proteina prionica (PrP)

umana si trova sul cromosoma 20. Il gene è attivo, e produce quindi la proteina, nel tessuto nervoso ed in altri tessuti, ad esempio quelli epiteliali. La proteina è presente sia nei tessuti normali che in quelli infetti, e la funzione che la proteina svolge normalmente non è nota; si stanno accumulando prove che la proteina può legare il rame. La PrP è associata alla membrana cellulare, mediante una porzione che si ancora al Glicosil-Fosfatidil-Inositolo ("àncora" GPI). La catena di amminoacidi della proteina si avvolge su se stessa presentando una struttura tridimensionale in cui prevale la configurazione ad elica (**alfa elica**).

• La proteina PrP può assumere anche una alterata struttura tridimensionale, prevalentemente a foglietto (**foglietto beta**), che conferisce alla proteina la peculiare **capacità di legarsi alle proteine in configurazione normale, provocandone la conversione nella forma alterata**. Le proteine così alterate si legano a loro volta ad altre proteine di struttura normale perturbandone la struttura, e così via. Questo meccanismo permette alla proteina di **"moltiplicarsi" in assenza di un genoma**, perché si avvia una reazione a catena che porta all'aumento progressivo della forma alterata (chiamata anche PrP-res, perché più resistente all'azione dell'enzima proteasi K) a spese della forma normale (chiamata anche PrP-sen, in quanto sensibile alla proteasi K).

L'insorgenza di alcune **malattie neurodegenerative nei mammiferi** è stata correlata a modificazione della proteina prionica cellulare. Queste malattie comprendono lo scrapie nella pecora, la BSE (encefalopatia spongiforme bovina) nei bovini, nota come "malattia della mucca pazza", e il morbo Creutzfeldt-Jakob (CJD) nell'uomo.

La modificazione della PrP ha origine con differenti meccanismi nelle diverse forme della malattia:

1) forma **infettiva**, per passaggio della PrP alterata da un individuo all'altro nell'ambito di una stessa specie (come ad es. nel kuru, una malattia da prioni associata al cannibalismo) oppure con trasmissione da una specie all'altra. Un fattore importante che controlla la suscettibilità di una specie a contrarre una encefalopatia spongiforme di un'altra specie è la somiglianza tra le proteine delle due specie: maggiore è questa somiglianza, maggiore è l'efficienza con cui la proteina PrP-res di una specie induce in provetta la conversione della PrP-sen di un'altra specie in PrP-res. L'efficienza della conversione in provetta è correlata alla trasmissibilità della malattia tra specie diverse.

2) Forma **sporadica**, insorta in individui senza storia familiare per la malattia; può essere dovuta a mutazione spontanea del gene per PrP in alcune cellule somatiche, o a una conversione spontanea della PrP nella forma alterata; in ogni caso, i casi si presentano in soggetti che hanno varianti del gene per la PrP che li rendono più suscettibili alla malattia.

3) Forma **genetica**, per una mutazione del gene per PrP ereditata dai genitori (eredità dominante: è sufficiente che uno dei due genitori possieda e trasmetta il gene mutato).

4) Di recente, è emersa una nuova forma della CJD, denominata "**nuova variante** della malattia di Creutzfeldt-Jakob" (**vCJD**). Questa malattia è apparsa nel marzo 1996, 10 anni dopo l'epidemia di BSE diffusa nelle mandrie bovine della Gran Bretagna (1986). Il tempo di incubazione è stimato in 10-30 anni. La vCJD è una malattia umana neurodegenerativa rara e fatale. Come la CJD, è una encefalopatia spongiforme trasmissibile, che si aggiunge alle forme di CJD già note: sporadica (85-90%), familiare (5-10%) e iatrogena (5%). Una descrizione della "Variant Creutzfeldt-Jakob disease (vCJD)" si trova sul sito dell'Organizzazione Mondiale della Sanità (http://www.who.int/mediacentre/factsheets/fs180/en/index.html).

In contrasto con le forme tradizionali di CJD, la vCJD colpisce pazienti più giovani (in media 29 anni, invece di 65), ha una durata relativamente più lunga (media di 14 mesi invece di 4,5) ed è probabilmente da collegare alla *esposizione, probabilmente attraverso l'alimentazione, all'agente della BSE dei bovini*.

Statistiche aggiornate sulla incidenza in Gran Bretagna della vCJD:
http://www.cjd.ed.ac.uk/data.html