



GHISLIERI
CENTRO PER LA
COMUNICAZIONE
E LA RICERCA

Progetto “Progressi in Biologia e Medicina”

17° Corso di formazione avanzata

**Epigenetica: dall’ereditarietà
transgenerazionale alla malattia**

23-25 maggio 2018, Collegio Ghislieri, Pavia

A cura di Carlo Alberto Redi

17° Corso di formazione avanzata

**Epigenetica: dall'ereditarietà
transgenerazionale alla malattia**



GHISLIERI
CENTRO PER LA
COMUNICAZIONE
E LA RICERCA

Progetto “Progressi in Biologia e Medicina”

17° Corso di formazione avanzata

**Epigenetica: dall’ereditarietà
transgenerazionale alla malattia**

23-25 maggio 2018, Collegio Ghislieri, Pavia

A cura di Carlo Alberto Redi

© Copyright 2018  EDIMES
Edizioni Internazionali - Pavia

Edizioni Internazionali srl
Divisione EDIMES - Edizioni Medico-Scientifiche - Pavia

Via Riviera, 39 - 27100 Pavia
Tel. 0382526253 - Fax 0382423120
E-mail: edint.edimes@tin.it

Tutti i diritti sono riservati.
Nessuna parte può essere riprodotta in alcun modo
(compresi i microfilm e le copie fotostatiche)
senza il permesso scritto dell'editore.

Indice

Prefazione	pag. VII
<i>CarloAlberto Redi</i>	
1. Epigenetica della trasmissione intergenerazionale dello svantaggio (o del vantaggio): no Lamarck	» 1
<i>CarloAlberto Redi</i>	
2. Il sociale si fa biologico: genomica sociale	» 8
<i>CarloAlberto Redi</i>	
3. Riprogrammazione epigenetica delle cellule germinali	» 18
<i>Manuela Monti</i>	
4. Pathways to neurodegeneration: beyond the genome	» 25
<i>Enza Maria Valente</i>	
5. Epigenomica e malattie neurodegenerative	» 29
<i>Cristina Cereda</i>	
6. Epigenetics, obesity and diabetes	» 42
<i>Lucia Migliore</i>	
7. Epigenetic mechanisms in early mammalian development	» 46
<i>Maria Elena Torres-Padilla</i>	
8. La natura epigenetica dei centromeri di mammifero	» 58
<i>Elena Raimondi</i>	
9. Staminalità e patogenesi	» 64
<i>Geppino Falco</i>	
10. The role of RNA in epigenetic regulation of gene expression	» 69
<i>Giuseppe Biamonti</i>	
11. Epigenetica e riprogrammazione cellulare	» 74
<i>Ileana Zucchi</i>	
12. Metodologie di indagine molecolare e strutturale applicate all'identificazione delle modificazioni post-traduzionali	» 82
<i>Federico Forneris</i>	

13. Malattie genetiche congenite da alterazioni dell'imprinting	»	90
<i>Orsetta Zuffardi</i>		
14. Gene therapy returns to centre stage	»	94
<i>Luigi Naldini</i>		
15. The epigenetic dimension of cancer: between development and plasticity	»	120
<i>Pietro Lo Riso</i>		
16. Experimental epidemiology: neurodevelopmental epigenetics across human genetic variability	»	127
<i>Nicolò Caporale</i>		
17. Non di solo DNA	»	134
<i>Giuseppe Testa</i>		

Prefazione

L'identità fenotipica delle cellule che compongono il nostro organismo è determinata dall'epigenoma: questo è ormai un dato acquisito. Quali siano i meccanismi e le architetture funzionali dell'insieme delle modificazioni chimiche, reversibili ed ereditabili, che controllano l'attività dei geni è oggi oggetto di studio a vari livelli, non solo quelli strettamente biomedici. Nel corso dello sviluppo embrionale vari modificatori epigenetici responsabili dello stabilirsi, mantenersi e rimuoversi delle modificazioni (la metilazione del DNA, il codice istonico, l'accessibilità della cromatina a diversi enzimi, l'organizzazione tridimensionale del genoma, il ruolo dei piccoli RNA non codificanti, per citarne alcuni) ne possono regolare una stabile ereditarietà o alterarsi anche in base alla presenza di xenobionti. Il corso di quest'anno è dedicato allo studio dei meccanismi che trasducono al DNA, alle cellule, ai tessuti, agli organi, a tutto il nostro corpo (compresa la nostra mente), i fattori chimici, fisici, sociali e culturali che definiscono l'ambiente nel quale si sviluppa l'intera storia del ciclo vitale di un individuo; in altre parole è dedicato allo studio dell'epigenomica ed all'analisi dei meccanismi epigenetici che controllano il regolare sviluppo ontogenetico. Di conseguenza, parte rilevante del corso è dedicata allo studio delle basi epigenetiche di diverse patologie, con una necessaria introduzione dedicata allo studio epigenetico delle cellule germinali. Tra i meccanismi causativi delle malattie non-trasmissibili (malattie cardio-vascolari, cancro, patologie croniche dei polmoni e diabete) un ruolo di rilievo è giocato dalle modificazioni epigenetiche di cellule germinali e cellule somatiche capaci di deregolarne l'espressione genica anche in termini di ereditarietà transgenerazionale: ecco dunque che si realizza una transizione sociobiologica del benessere o del malessere che le condizioni di natura e di cultura in cui si sviluppa e vive un individuo sono in grado di determinare. Ed in una situazione che si fa ricorsiva, le nuove impronte epigenetiche determinano un vantaggio o uno svantaggio che rincorre, influenzandosi reciprocamente, le condizioni ambientali in una relazione circolare che è in grado di determinare multimorbidità: basterà ricordare la compromessa spermatogenesi degli obesi capace di trasmettere alle generazioni future suscettibilità a diverse malattie croniche e alla stessa obesità. Ecco dunque che dal profilo strettamente biomedico lo studio dell'instaurarsi e della trasmissione intergenerazionale delle modificazioni epigenetiche diviene di interesse della epidemiologia, della sociologia, della economia e ... di molte altre

discipline, non ultime del diritto e della filosofia chiamando in causa il decisore politico affinché sia reso edotto e consapevole che il proprio agire non può essere alieno a queste conoscenze.

Nell'insieme lo sforzo didattico di tutti i relatori (ai quali va uno speciale ringraziamento) ha permesso di produrre il volume degli atti che certamente aiuterà a rafforzare quanto in aula verrà presentato e discusso.

Un grazie affettuoso all'Amministrazione del Collegio che organizza i nostri corsi ed a tutte le persone che con grande competenza professionale (e pazienza!) ne permettono la realizzazione.

CarloAlberto Redi

Epigenetica della trasmissione intergenerazionale dello svantaggio (o del vantaggio): no Lamarck

CarloAlberto Redi

Dipartimento Biologia e Biotecnologie “Lazzaro Spallanzani”, Università di Pavia

Di recente è stato dimostrato che la compromessa spermatogenesi degli obesi è in grado di trasmettere alle generazioni future suscettibilità a diverse malattie croniche ed alla stessa obesità (Fullston et al., 2013): questa trasmissione intergenerazionale si attua grazie agli spermatozoi danneggiati e portatori di marcature epigenetiche (prodotte dai contesti ambientali, stile di vita, dieta, etc.) a livello di diversi geni implicati nella suscettibilità a queste malattie (insulino resistenza, tolleranza al glucosio, etc.) che si presenteranno nella progenie. È noto che il diabete paterno di tipo 2 è in grado di influenzare i tratti metabolici della progenie e probabilmente ciò si attua sia via spermatozoi portatori di alterati stati di metilazione del DNA (Wei et al., 2014a), sia portatori di piccole molecole di RNA non-codificante (micro-RNA, molecole capaci di interferire, regolandolo, con il processo di traduzione dell'informazione genetica da DNA a proteine) così come è stato osservato negli spermatozoi di uomini obesi (Donkin et al., 2016). Sebbene un classico dogma in Biologia ricordi che il genoma degli spermatozoi è silente sotto il profilo della trascrizione del DNA, è oggi noto che gli spermatozoi veicolano una quantità di piccole molecole di RNA non codificante proteine; inoltre vi sono sempre più chiare evidenze che la composizione di questi micro-RNA negli spermatozoi sia determinata da fattori ambientali (Gapp et al., 2014) ed in particolare quelli che si originano in seguito a diete alterate e sbilanciate sono in grado di determinare caratteristiche ereditabili (Grandjean et al., 2015; Donkin et al., 2016). Le piccole molecole di micro-RNA sono oggi al centro di intensi programmi di ricerca poiché sono candidati ideali per lo stabilirsi e per il trasmettersi della “memoria” delle condizioni ambientali entro le quali si sono sviluppate le cellule germinali; memoria che si attua con il silenziamento o l'iperattività di geni la cui espressione può essere regolata differenzialmente da entrambi i processi: sia da quello di metilazione delle citosine (*imprinting*) sia da quello dell'interferenza esercitata dai micro-RNA sulla loro regolare espressione (*RNA interference*).

Non solo una condizione di obesità ma anche di malnutrizione (denutrizione) nel corso della vita *in utero* si dimostra capace di alterare la futura composizione degli spermatozoi (degli individui che hanno sofferto quella condizione *in utero*) e di essere così trasmessa per diverse generazioni (Radford et al., 2014). Tutti i

processi biologici sottesi allo sviluppo embrionale sono “complessi” e però quelli della gametogenesi si rivelano di una particolare sensibilità alla presenza di circostanze “sfavorevoli” di qualsivoglia natura (dalle alterazioni delle condizioni chimico-fisiche quali la temperatura, pH, etc. alla presenza/assenza di svariati elementi e sostanze quali ormoni, inquinanti alimentari, metalli, etc.) capaci di alterare il processo di cito-differenziazione delle cellule germinali. Spermatozoi e cellule uovo si formano nel corso di un periodo temporale estremamente lungo (dalla vita embrionale alla adolescenza sino alla senescenza, menopausa per le femmine; si veda il contributo di Manuela Monti in questo volume: “Riprogrammazione epigenetica delle cellule germinali”) e dunque costantemente soggetto ad alterazioni delle normali condizioni richieste per un corretto sviluppo; inoltre per raggiungere uno stadio di maturità funzionale vanno incontro anche a complessi processi cito-differenziativi che da una cellula rotondeggiante li porterà ad acquisire una forma definitiva necessaria per il buon andamento della fecondazione. A complicare questi eventi morfo-genetici vi sono i processi citologici e molecolari della riduzione meiotica del patrimonio genetico di ciascuna cellula germinale. E dunque non stupisce che le condizioni ambientali entro le quali si sviluppano i processi gametogenetici siano in grado di stabilire la costituzione citologica e molecolare delle cellule uovo e degli spermatozoi: durante le prime fasi dello sviluppo embrionale le cellule germinali primordiali migrano dal luogo della loro origine sino a quello ove si costituisce l'abbozzo delle gonadi (ovario e testicolo; si veda ancora il contributo di Manuela Monti in questo volume), contraggono rapporti citologici con le cellule somatiche delle gonadi: nella femmina iniziano a proliferare, aumentano di numero, iniziano il processo meiotico, si bloccano in una precisa fase (in termini gergali, diplotene) e solo con il menarca porteranno a termine il processo meiotico con una precisa cadenza mensile; nel maschio, una volta raggiunta la sede ove si sviluppa il testicolo si organizzano i tubuli seminiferi e le cellule primordiali (spermatogonio) si bloccano iniziando e portando a termine il processo meiotico solo alla pubertà. Condizioni avverse che si realizzino nel corpo materno (obesità, malnutrizione, denutrizione, etc.) sono intuitivamente in grado di perturbare il corretto svolgimento di una così intricata architettura di relazioni fisiologiche quali quelle che portano alla formazione di cellule germinali mature e così creano le premesse per una trasmissione alle generazioni successive di alterati programmi genetici di sviluppo (Radford et al., 2014). Ciò che è bene precisare, è che si realizza un'alterata riprogrammazione genetica delle cellule germinali. Nel corso del regolare processo di gametogenesi vengono cancellate tutte le marcature epigenetiche (essenzialmente de-metilazione del DNA) acquisite dalle cellule germinali e ne vengono stabilite di nuove, quelle che si sono andate evolutivamente selezionando per assicurare un regolare processo di sviluppo dell'embrione. Se il programma di cancellazione delle marcature epigenetiche (essenzialmente metilazione delle citosine, acetilazione di proteine istoniche) non si realizza, queste ultime vengono trasmesse alla progenie, l'embrione eredita quelle imposte dalle condizioni ambientali alle cellule germinali dei genitori, l'obesità e la suscettibilità alle malattie croniche (cardiocircolatorie, respiratorie, insulino resistenza, etc.) passano di generazione in generazione per almeno una,

due, tre generazioni e forse più. Questo è il quadro che emerge dagli studi epidemiologici e da quelli sperimentali su modelli animali che dimostrano la persistenza di regioni del DNA differenzialmente metilate nel DNA degli spermatozoi di maschi obesi; regioni del genoma capaci di resistere alla riprogrammazione e di determinare malattie nella generazione seguente (Kobayashi et al., 2012). Questi dati assumono un rilievo particolare quando si pensi alla riproduzione umana in linea generale poiché sanciscono la possibilità di trasmettere marcature epigenetiche, acquisite dalle cellule germinali a causa delle condizioni socio-economiche vissute dai portatori, alle generazioni successive: padri in condizioni di pre-diabete aumentano significativamente la probabilità di contrarre il diabete ai propri figli per la trasmissione di alterati quadri genici di metilazione di citosine; i quadri di metilazione degli spermatozoi di padri in pre-diabete possono essere trasmessi per ben due generazioni alle cellule delle isole pancreatiche favorendo così lo sviluppo del diabete nella progenie (Wei et al., 2014b). Non solo la metilazione delle citosine di alcune sequenze geniche è alterata in queste condizioni ma anche le proteine istoniche che avvolgono il DNA e la costituzione dei micro-RNA non codificanti. La nota positiva in questo scenario è che l'identificazione delle marcature epigenetiche imposte dalle condizioni ambientali alle cellule germinali ne dovrebbe permettere l'identificazione anche in cellule quali i globuli polari della cellula uovo e negli spermatozoi stessi, permettendo così di sviluppare metodi diagnostici predittivi dell'instaurarsi di quelle condizioni negative nella progenie grazie ad analisi genetiche precedenti la fecondazione *in vitro*.

Uno dei fattori, forse in alcuni contesti socio-economici IL fattore, più rilevante ai fini di una corretta gametogenesi è costituito dalla dieta. Ad essere dannosa per la oogenesi (e la spermatogenesi) non sono solo la sovra-alimentazione ed una dieta particolarmente ricca di zuccheri, sale e comunque una dieta ipercalorica: va infatti ricordata l'alimentazione carente di nutrienti essenziali sofferta da miliardi di persone, una condizione drammatica. I processi di oogenesi e di spermatogenesi risultano particolarmente sensibili alla carenza di fattori provitaminici, di vitamine e oligoelementi. Una gran parte della popolazione del pianeta Terra si trova dinanzi a due scenari altrettanto dannosi per la salute: se da un lato si riscontra la produzione di "oociti obesi" dall'altro quello della produzione di oociti non in grado di sostenere lo sviluppo embrionale, in seguito alla carenza di vitamine, minerali, micro- e macro-nutrienti essenziali. Può instaurarsi una condizione di malnutrizione anche in presenza di regolare accesso al cibo: è questo il caso di coloro che seguono regimi alimentari errati (diete sbilanciate, etc.). Un bambino malnutrito è condannato ad un minore sviluppo fisico e mentale nel corso dell'infanzia. In oltre 147 milioni di bambini in età pre-scolare nei paesi in via di sviluppo si riscontrano ritardi nella crescita. La mancanza di iodio è la prima causa di ritardo mentale e di danni cerebrali. La denutrizione incide negativamente sul rendimento scolastico ed in una reazione a catena ciò determina una incapacità a generare reddito una volta adulti. Inoltre, una madre denutrita ha più probabilità di dare alla luce bambini sottopeso. I primi due anni di vita del bambino sono fondamentali per prevenire la denutrizione infantile, causa, in gran parte, di danni irreversibili. Le ultime statistiche della FAO contano 795 milioni di persone che soffrono la

fame nel mondo, di cui il 98% vive nei paesi in via di sviluppo. La distribuzione nei continenti è la seguente: 511,7 milioni in Asia; 232,5 milioni in Africa; 34,3 milioni in America Latina e Caraibi; 14,7 milioni nei Paesi sviluppati. Si stima che siano 200 milioni i bambini che, nei paesi in via di sviluppo, soffrono di una qualche forma di malnutrizione (la malnutrizione dei bambini, UNICEF, 2012). Purtroppo, spesso la fame è ereditata: sono ben 17 milioni i bambini che ogni anno nascono sottopeso, a causa di un'insufficiente alimentazione materna, prima e durante la gravidanza. Le donne sono il primo produttore di cibo al mondo. Eppure, tradizioni culturali e strutture sociali spesso inducono le donne ad essere maggiormente colpite dalla fame e dalla povertà rispetto agli uomini. Una madre sottopeso a causa di un'alimentazione inadeguata ha più probabilità di dare alla luce un bambino sottopeso rispetto ad una madre ben nutrita. Circa il 50 per cento delle donne in stato di gravidanza nei paesi in via di sviluppo soffre di mancanza di ferro (UNICEF, 2012; Hanson et al., 2015) e ciò significa che 315.000 donne muoiono ogni anno per emorragie durante il parto. Per questo, le donne, specialmente se in gravidanza o in fase di allattamento, necessitano di alimenti specifici o in maggior quantità. La carenza di micronutrienti, vitamine e minerali, colpisce quasi due miliardi di persone nel mondo. In base alle statistiche dell'Organizzazione Mondiale della Sanità, la mancanza di ferro, vitamina A e zinco è fra le prime dieci cause di morte per malattia nei paesi in via di sviluppo. La mancanza di ferro è la forma di malnutrizione più diffusa, che colpisce miliardi di persone nel mondo. La mancanza di ferro aumenta il rischio di malattie nei bambini, ostacola lo sviluppo cognitivo delle persone e, in ultima analisi, ha effetti negativi sulla produttività della popolazione di un paese. La mancanza di vitamina A è la prima causa di cecità fra i bambini dei paesi in via di sviluppo inoltre, può aumentare il rischio di morte per diarrea, morbillo e malaria. La mancanza di iodio colpisce 780 milioni di persone nel mondo e circa 20 milioni di bambini nascono con deficit mentali a causa di una carenza di iodio delle madri durante la gestazione. La mancanza di zinco influisce negativamente sulla crescita e indebolisce l'apparato immunitario nei bambini in età pre-scolare: è causa di morte per 800.000 bambini, a livello mondiale, ogni anno. Come facilmente immaginabile anche una nutrizione povera durante il concepimento, la gravidanza e l'allattamento può avere effetti gravi sulla corretta crescita del feto, provocando danni che si estendono per tutta la vita del bambino. Uno tra questi è lo *stunting*, ovvero un processo di ritardo o blocco della crescita che inizia in utero e continua per i primi due anni di vita del bambino. In Stati come lo Zambia o l'India, dove la percentuale di malnutrizione è molto alta e le donne hanno poco accesso alle cure prenatali, i casi di *stunting* nei bambini sono altissimi, così come i dati relativi alla mortalità infantile. In Zambia, nel 2014 ad esempio, si sono registrati circa 87 morti ogni 1.000 nati vivi. Ci sono poi le conseguenze di lungo periodo, alla base di un grave circolo vizioso. I problemi nello sviluppo cognitivo, frequenti nei bambini affetti da *stunting*, compromettono l'istruzione dei piccoli in età scolare e generano una futura perdita di forza lavoro, di capacità produttiva, perpetuando uno stato di povertà in genere già molto diffuso in Paesi che vivono queste condizioni: si stima che i costi legati alla malnutrizione ammontino, globalmente, a circa 30 miliardi di dollari

l'anno. Un nutrito numero di ginecologi ed ostetrici, riuniti sotto l'egida di una federazione internazionale attenta ai legami tra nutrizione, sviluppo e conseguenze per la salute e la economia, ha preparato un interessante rapporto (Hanson et al., 2015) dal titolo "*Think Nutrition First*" ove sono dettagliate tutte queste ultime considerazioni e statistiche.

Da tutte le evidenze sperimentali su modelli animali e dalle evidenze epidemiologiche in umano risulta dunque chiaro che le condizioni ambientali entro le quali si snoda la storia del ciclo vitale dei genitori è in grado di influire sensibilmente sulle traiettorie di salute incontrate dalla progenie grazie ai cambiamenti epigenetici registrati dalle cellule germinali. La capacità di trasmettere alle generazioni successive le acquisite (dal contesto ambientale) caratteristiche genomiche non deve però lasciar intendere che ci si trovi dinnanzi ad una forma di ereditarietà di tipo lamarchiano, in altri termini, di trovarsi dinnanzi ad una ereditarietà dei caratteri acquisiti e ad una smentita clamorosa dei paradigmi darwiniani. È qui necessaria una precisa puntualizzazione: la trasmissione dell'acquisita suscettibilità all'obesità, ad esempio, è del tutto temporanea e si trasmette per una o al massimo poche generazioni ed è del tutto reversibile nel momento in cui cambiano i contesti ambientali (lo svantaggio sociale dei disereditati). La metafora del collo delle giraffe può essere utile: il collo delle giraffe si allunga a mano a mano che le chiome degli alberi si innalzano poichè la selezione naturale elimina all'interno della popolazione di giraffe quelle con collo meno lungo (incapaci di competere per raggiungere le chiome e alimentarsi, quindi riprodursi, quindi trasmettere alla progenie la propria costituzione genetica per colli meno lunghi di quelli di altre giraffe). La trasmissione dell'allungamento del collo alla progenie si realizza in uno scenario darwiniano di evoluzione per selezione naturale contro gli individui che portano una costituzione genetica tale da produrre una statura insufficiente (più bassa) di altri che portano una costituzione genetica capace di determinare un collo più lungo (maggiore altezza): nel passaggio da una generazione alle successive nella popolazione di giraffe, si realizza una variazione delle frequenze alleliche a favore di quegli alleli che determinano un lungo collo; in altri termini, i piccoletti non riescono a cibarsi, a riprodursi e quindi a passare alle generazioni successive la costituzione per collo corto, i cui alleli vengono eliminati dalla popolazione.

Suggeriamo fortemente di approfondire questi argomenti visitando un sito web in lingua italiana e di facile lettura, (<http://epigenome.eu/it/4,14,0>), sebbene di rigorosa presentazione: "*Epigenome Network of Excellence*" (una rete di ricerca europea che si dedica al supporto di iniziative scientifiche di alto livello nel campo dell'epigenetica) per capire meglio in che modo l'epigenetica conferisca "forma" alla vita. Riassumendo, come ben noto, ciascuna cellula in un organismo contiene la medesima quantità di DNA ma solo uno specifico gruppo di geni è espresso in ciascun tipo cellulare. Nel corso dello sviluppo embrionale, si realizza la differenziazione dei diversi tipi cellulari che vede nelle modificazioni delle componenti proteiche del genoma uno dei principali meccanismi capaci di assicurare un fine e articolato processo di regolazione dell'espressione genica silenziando o attivando specifici geni. È così che metilazione, acetilazione e fosforilazione di specifici

aminoacidi agiscono nello specificare lo stato trascrizionale dei geni nel corso dello sviluppo (in termini tecnici, il classico esempio è la metilazione della lisina in posizione 9 o 27 dell'istone H3 per reprimere l'espressione di un gene) ed assicurano meccanismi di mantenimento a lungo termine (per più generazioni) della memoria epigenetica imposta dalle condizioni ambientali (De e Kassis, 2017). Tuttavia, i meccanismi molecolari sui quali si basa la trasmissione intergenerazionale dell'ereditarietà epigenetica restano ancora poco conosciuti (Miska e Ferguson-Smith, 2016; Sharma et al., 2017) nonostante il fatto che si conoscano bene i tempi ed i modi della riprogrammazione epigenetica che avviene nelle cellule germinali, nel corso del corretto sviluppo embrionale, ove di fatto viene "cancellato" ogni modificazione acquisita (Morgan et al., 2005) e ne viene re-imposta un'altra funzionale allo sviluppo embrionale. Evidentemente, per effetto dell'esposizione a fattori ambientali negativi così come brevemente sopra illustrato, qualcuno di questi meccanismi si inceppa e si deregola portando al passaggio generazionale delle modificazioni acquisite dalle cellule germinali. La metilazione del DNA, le modificazioni istoniche (i.e. acetilazione) ed i piccoli RNA a capacità regolatoria sono i più probabili mediatori della transizione socio-biologica costituendosi quali portatori di informazione epigenetica capace di passare da una generazione all'altra: senza scomodare Lamarck e per la tranquillità di Darwin!

Bibliografia

1. De S, Kassis J. Passing epigenetic silence to the next generation, in "Science". 2017; 356: 28-29.
2. Donkin I, et al. Obesity and Bariatric Surgery Drive Epigenetic Variation of Spermatozoa in Humans, in "Cell Metabolism". 2016; 23: 369-378.
3. Fullston T, et al. Paternal obesity initiates metabolic disturbances in two generations of mice with incomplete penetrance to the F2 generation and alters the transcriptional profile of testis and sperm microRNA content, in "FASEB Journal". 2013; 27: 4226-4243.
4. Gapp, K. et al. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat. Neurosci.* 2014; 17: 667-669.
5. Grandjean V, et al. RNA-mediated paternal heredity of diet-induced obesity and metabolic disorders, in "Scientific Reports". 2015; 5: Article number: 18193.
6. Hanson M, et al. The International Federation of Gynecology and Obstetrics (FIGO) recommendations on adolescent, preconception, and maternal nutrition: "Think Nutrition First", in "International Journal of Gynecology and Obstetrics". 2015; 131: S213-253.
7. Kobayashi H, et al. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks, in "PLoS Genet". 2012; 8: e1002440.
8. Miska E, Ferguson-Smith A. Transgenerational inheritance: Models and mechanisms of non-DNA sequence-based inheritance, in "Science". 2016; 354: 59-63.

9. Morgan H, et al. Epigenetic reprogramming in mammals, in "Human Molecular Genetics". 2005; 14: R47-R58.
10. Radford E, et al. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism, in "Science". 2014; 345: 785-793.
11. Sharma U, Rando O. Metabolic Inputs into the Epigenome, in "Cell Metabolism". 2017; 25: 544-558.
12. Wei Y, et al. Paternally induced transgenerational inheritance of susceptibility to diabetes in mammals, in "Proc. Natl. Acad. Sci. USA". 2014; 111: 1873-1878.
13. Wei Y, Schatten H, Sun Q. Environmental epigenetic inheritance through gametes and implications for human reproduction, in "Human Reproduction Update". 2014; 21: 1-15.

Il sociale si fa biologico: genomica sociale

CarloAlberto Redi

Dipartimento Biologia e Biotecnologie “Lazzaro Spallanzani”, Università di Pavia

La “transizione socio-biologica” è stata preconizzata da Rudolf Virchow (1821-1902) nel corso di una epidemia di tifo nel 1848 con la sua visione della patologia quale strumento per capire il legame tra condizioni sociali e malattie: la terapia più efficace contro le malattie è quella in grado di assicurare un buon grado di educazione, istruzione, prosperità economica, assistenza per l’infanzia e libertà. Per queste affermazioni il governo prussiano lo isolò e non ricevette il premio Nobel per la fisiologia o la medicina nel 1902. Dalla intuizione di Virchow ad oggi si sono accumulate una grande quantità di evidenze scientifiche che provano come il contesto sociale nel quale si sviluppa la storia del ciclo vitale degli individui sia capace di influenzare molti processi biologici e così il sociale si “incarna” nel biologico e si trasmette da una generazione all’altra. Il modello del “periodo critico” (Underwood, 2014) spiega le differenze in salute osservate nelle varie classi sociali in dipendenza dell’esposizione differenziale di ciascuna fase dello sviluppo biologico a specifici fattori negativi: deprivazione alimentare ed emotiva, malattie, comportamenti rischiosi quali fumo, diete sbilanciate, uso di droghe, etc etc. Tra i grandi ricercatori che hanno contribuito in modo significativo all’avanzamento delle conoscenze scientifiche nel campo della epidemiologia sociale spicca certamente la figura di Sir Michael Marmot con il compendio di risultati, ottenuti in una vita di ricerche condotte in USA, Australia, Brasile, Finlandia, Cuba, Canada, ed altri paesi ancora, raccolti nel monumentale *“The Status Syndrome: How Social Standing Affects Our Health and Longevity”* dell’anno 2004; il titolo già anticipa il messaggio portante della riflessione di Sir Marmot: la posizione occupata nel gradiente sociale colpisce direttamente la nostra salute. Al termine di accurate ricerche e riflessioni (già premiate nel 2004 con il prestigioso premio Balzan), Sir Marmot rivolge delle raccomandazioni a chi ha il dovere di predisporre programmi di politiche sociali, raccomandazioni che ricordano quanto scrive Virchow ai governanti, e divenute famose come le 6 raccomandazioni di Marmot. Infatti, per prevenire e guarire le malattie è necessario assicurare ai più svantaggiati (ma il richiamo vale per la società civile tutta):

- più assistenza per l’infanzia,
- migliore istruzione,
- più occupazione,
- salario minimo garantito per tutti,

- comunità più sane e sostenibili, più case accessibili a tutti,
- determinazione sociale ad eliminare fumo e consumo di alcol.

Raccomandazioni che hanno ricevuto sostegno da tanti studiosi ma assai raramente considerate dai decisori politici. Così Grace Budrys nel 2010 con il suo *“Unequal health: how inequality contributes to health or illness”* rinforza in modo chiarissimo quanto sostenuto ormai da tutta la epidemiologia sociale: più in basso nella gerarchia socioeconomica una persona viene a trovarsi, peggiore sarà la sua condizione di salute e questa condizione passa da una generazione all'altra, il sociale si fa biologico. Non vi sono solo dati di epidemiologia a sostenere questa tesi, anche ricerche sperimentali condotte su diverse specie di Primati vengono presentate e ne emerge che la posizione occupata nella scala gerarchica influenza potentemente la salute anche nei nostri lontani parenti. E la condizione di benessere o mal-essere si determina a partire da stadi molto precoci del nostro sviluppo, in utero; in poche parole, dipende dal gradino della scala sociale occupato dai nostri genitori. Sir Marmot si è poi concentrato sui fenomeni inter-generazionali che riguardano in particolare il comportamento della madre e che sono capaci di determinare il nostro stato di salute via trasmissioni inter-generazionali non-genetiche. I risultati di queste ricerche sono pubblicati in un interessante libro dell'anno 2015 dal chiaro titolo *“La salute disuguale. La sfida di un mondo ingiusto (Il Pensiero Scientifico Editore, Roma, 2016)* e sono stati presentati dallo stesso Sir Marmot nel corso di un interessante convegno alla Commissione Sanità del Senato (svolto in sala Zuccari di Palazzo Giustiniani il martedì 24 gennaio 2017) dal titolo *“Disuguaglianze di salute: quanto pesano? Si possono modificare?”* È possibile scaricare l'intera presentazione (al sito <http://www.epicentro.iss.it/temi/politichesanitarie/pdf/Senate%20Committee%20Rome%20Jan%202017.pdf>) corredata da numerosissimi grafici e tabelle. Vi sono immagini in grado di comunicare in modo drammatico la relazione tra stato di censo, stato di salute ed aspettative di vita. Alcune statistiche presentano numeri difficili da credere se la fonte non fosse tanto autorevole. Il lettore è invitato a scaricare l'intera presentazione e solo a titolo di esempio riporto, e stiamo parlando di Inghilterra, i dati su “stili di vita e nutrizione” (non vi è bisogno di traduzione):

- Only 18% of people have one or more meals a day at their table;
- Nearly two thirds of people eat at their table less than once a week;
- 30% of households use their table for meals barely a few times a year;
- 3% have no table.

Commoventi le immagini finali con affermazioni (*“Social injustice is killing on a grand scale”*) che dovrebbero far riflettere tutti coloro che hanno responsabilità politica nella gestione del benessere di una società poiché basate su evidenze scientifiche. Venerdì 2 giugno 2017 nel corso del festival di Trento dedicato alla salute disuguale Sir Marmot (con Luca De Fiore) ha ancora una volta delineato in modo magistrale questi concetti parlando di *“Vita e morte nella scala sociale”*; vi sono drammatiche differenze nello stato di salute tra diversi paesi e all'interno di un singolo paese e non si tratta solo e semplicemente di povertà e ricchezza.

Un povero di Glasgow è ricco in confronto ad un indiano con un reddito medio ma il povero di Glasgow ha una aspettativa di vita inferiore di ben 8 anni rispetto all'indiano; è molto probabile che l'indiano morirà di patologie infettive legate alla sua povertà mentre il povero di Glasgow di morte violenta, suicidio, attacco cardiaco, tutte cause di morte legate alla versione di svantaggio tipica della sua società. In tutti i paesi, le persone con un relativo svantaggio sociale, soffrono di uno svantaggio in salute. Più è alto lo stato sociale di un individuo, migliore è la sua salute, vivrà più a lungo ed avrà una vecchiaia più in salute. Quello che rende queste disuguaglianze di salute ingiuste è la chiara evidenza che i governanti sanno bene cosa potrebbe almeno far diminuire questa ingiustizia: i dati accumulati dalla comunità scientifica grazie a studi a livello internazionale indicano che la presa in carico delle raccomandazioni di Sir Marmot avrebbe la forza di cambiare radicalmente il modo in cui si pensa alla salute ed anche alle società in cui viviamo. Per il lettore che desidera approfondire alcune di queste tematiche è di grande aiuto visitare il sito dell'istituto universitario ove Sir Marmot svolge le proprie ricerche (*University College London, Institute of Health Equity*, www.institute-ofhealthequity.org). I risultati delle ricerche di Sir Marmot non hanno solo messo in luce questi grandi portati dell'epidemiologia sociale ma hanno anche fornito dettagli di assoluto rilievo per la salute ed il benessere della vita quotidiana di milioni di cittadini. Dettagli che parrebbero a prima vista sottili disquisizioni sui casi considerati ma che in realtà svelano la presenza di fattori cruciali della realtà sociale che trascendono la mera condizione economica e di reddito, rivelano l'esistenza di fattori psicologici capaci di "incarnarsi", di farsi biologici. Si pensi ai primi studi condotti sui dipendenti statali inglesi (Whitehall I e Whitehall II) ove è in grado di dimostrare una diminuzione dell'aspettativa di vita in relazione alla posizione gerarchica occupata nello svolgimento della propria professione in base alla deregolazione di equilibri metabolici ed endocrine capaci di portare ad un generale peggioramento della salute con maggiore facilità a sviluppare malattie. La variazione di posizione gerarchica induce fluttuazioni del proprio sentirsi padroni delle condizioni di lavoro che a loro volta portano ad una diminuita integrazione sociale, perdita di controllo sul proprio destino e di autostima ed, a cascata, una miriade di comportamenti (cambiamenti nello stile di vita, alimentazione, fumo, alcool, depressione, *etc.*) che aumentano l'isolamento sociale a causa del ruolo di subalternità vissuto come mancato riconoscimento del proprio ruolo sociale e professionale: il grado di autodeterminazione si rivela così un fattore psico-sociale molto importante per la tutela della salute (Sherman et al., 2012).

Nel corso dello sviluppo della storia del ciclo vitale di un individuo (cellule germinali - embrione - feto - giovane - adulto - ((senescente)) - cellule germinali - embrione - ...) le cellule, i tessuti, gli organi sono esposti a diversi ambienti. Il termine ambiente va considerato nella più ampia accezione: per le cellule germinali è ambiente l'ovario o il testicolo, per l'embrione l'organismo materno, per i nuovi individui l'aria, l'acqua, la famiglia, la scuola, *etc.* La prima fase dello sviluppo, dal concepimento alla precoce vita intrauterina sino alla adolescenza è un periodo di crescita e sviluppo da una singola cellula uovo fecondata ad un individuo adulto. La seconda fase, all'incirca dopo i vent'anni, è un periodo di declino

dal massimo momento di crescita alla perdita di funzione, alle malattie ed alla morte. Il genoma (DNA) nelle diverse fasi dello sviluppo è esposto ad una varietà di agenti chimici e fisici (xenobionti); l'ambiente sociale (censo, famiglia, scuola, religione, cultura, etc.) ne influenza in modo determinante il grado di esposizione e la struttura sociale tende a veicolare continuità di vantaggi o svantaggi: sono ben noti sia l'arresto della crescita in altezza dovuto a deprivazioni emotive o nutrizionali degli infanti sia le marcate differenze in longevità, aspettativa di vita in buona salute e forma fisica in età avanzata in relazione alla classe sociale.

Le ricerche sono rivolte a chiarire i meccanismi attraverso i quali "il sociale entra nella pelle e si fa biologia"; del come la classe sociale entra nelle molecole, nelle cellule. Alcuni dei più robusti archivi di dati riguardano gli studi longitudinali effettuati su diverse malattie (si veda UK *Economic and Social Research Council* e USA *National Institute of Aging*). Classico lo studio sugli effetti della deprivazione sociale sulle traiettorie di salute nei pazienti affetti da fibrosi cistica: considerando che non vi sono differenze socio-economiche nella incidenza di fibrosi cistica, poiché la malattia ha una origine genetica, si possono evidenziare traiettorie sociali ben precise per le condizioni patologiche della malattia in dipendenza del grado di deprivazione sociale, accesso al sistema delle cure e stato occupazionale. I dati rivelano aspetti drammatici della transizione socio-biologica: i pazienti più agiati dimostrano una migliore funzione polmonare ed una minore colonizzazione da parte del batterio *Pseudomonas aeruginosa* (prima causa di gravi infezioni polmonari in questi pazienti), i trattamenti terapeutici differiscono fortemente in base al gruppo sociale e le ineguaglianze vanno sempre aumentando (opportunità di lavoro, severità della patologia, tempi di ospedalizzazione, etc.). Anche lo studio delle relazioni tra massa corporea e salute cardio-metabolica rivela la relazione tra peso (nelle diverse fasi dello sviluppo, da quelle gestazionali a quelle dell'adulto) e rischio coronarico con traiettorie di esito marcatamente diverse in base alla classe sociale. Ed ancora, lo studio della salute del cavo orale (denti, gengive e bocca) mostra come la relazione tra igiene orale, stato di salute e qualità della vita dipenda dal censo. Tra gli studi di maggior rilievo vi sono quelli sullo sviluppo delle cellule germinali, in particolare lo sviluppo della cellula uovo e delle prime fasi embrionali/fetali, in relazione ai disordini nutrizionali ed agli effetti negativi che direttamente e indirettamente vengono da questi esercitati sul metabolismo della cellula uovo (e dei suoi organelli, in primis i mitocondri) e sui processi citologici a cui va incontro dopo fecondazione determinando difetti nello sviluppo, malformazioni congenite e danni alla salute della progenie. Individui obesi tendono ad avere un numero inferiore di cellule germinali e di embrioni vitali; un eccesso di colesterolo o di sale è capace di diminuire significativamente il numero di follicoli ovarici e di determinare sterilità femminile; il rischio di sviluppare diabete per disagiate condizioni durante la gravidanza, povero sviluppo del feto, scarso peso alla nascita è chiaro, il rischio di diabete in uomini oltre i 60 anni dipende dal peso alla nascita: peso di circa 4 kg = rischio 1; peso di circa 2,5 kg = rischio 7-8 volte maggiore).

Anche le relazioni che gli individui contraggono evidenziano transizioni socio-biologiche: Ciascuno di noi vive in un contesto di interdipendenza da altri

individui e condivide influenze ambientali simili (reddito familiare, divisione dei lavori domestici, preferenze alimentari o di vacanze): è noto che la diagnosi di una grave malattia e l'angoscia ad essa associata hanno un impatto negativo sia sul paziente che sul coniuge o che la depressione di un compagno colpisce le facoltà cognitive dell'altro compagno.

Quanto il gradiente sociale sia deleterio per la salute emerge da studi ormai divenuti classici. Le disparità socio-economiche (oltre alla rapida urbanizzazione in corso a livello mondiale, le diete e l'invecchiamento della popolazione) contribuiscono in modo determinante alle prime cause di mortalità a livello mondiale: malattie cardio-vascolari, cancro, patologie croniche dei polmoni e diabete, causano in totale circa 40 milioni di morti all'anno. Occupano dunque un posto di rilievo nei meccanismi causativi di tutte quelle che vengono definite "malattie non trasmissibili" (*non-communicable diseases*), per definizione malattie non infettive e non trasmissibili tra le persone. Uno dei più recenti ed interessanti lavori della scuola di David Blane (*Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK*) entra nei dettagli del riscontro sociale in chiave di "operatore del farsi biologico", andando ad analizzare le specifiche patologie indotte dal gradiente sociale (McLean et al., 2014). È noto infatti che si riscontra una plurimorbilità in età giovanile negli individui che vivono in aree ad alta deprivazione socioeconomica ma ben poco si sa sulla tipologia delle plurimorbilità nelle diverse fasce di età in relazione allo stato socioeconomico. I dati sociosanitari di ben 1.272.685 adulti scozzesi sono stati scomposti e classificati in base all'età, socio-deprivazioni, patologie fisiche, mentali e fisico-mentali. È così emerso che le sole patologie fisiche (il 56% di tutte le patologie) sono le patologie più comuni per gli individui con più di 55 anni e non variano sensibilmente in relazione allo stato socioeconomico; le sole patologie mentali non sono frequenti (4% di tutte le patologie) mentre le patologie fisico-mentali rappresentano ben il 40% di tutte le patologie nel gruppo al di sotto dei 55 anni e, fatto di estremo interesse, dimostrano un significativo incremento (raddoppiandosi o triplicandosi) in ciascun sottogruppo di età tra gli individui più deprivati e quelli meno deprivati in relazione allo stato socio-economico. Nell'insieme i dati dimostrano che la multimorbilità aumenta in relazione alla fascia di età, passando dall'8% per i 25-34 anni al 76% per quelli di un'età superiore ai 75 anni. Inoltre, risulta chiaro che in tutte le fasce di età considerate la deprivazione socioeconomica correla con la prevalente presenza di depressione ed il maluso dei medicinali, ansietà, dispepsia, dolore, patologie coronariche, diabete per tutti i pazienti con multimorbilità.

Un originale contributo agli studi di epidemiologia sociale viene anche da un'altra figura particolare di studioso, Paolo Vineis (professore di epidemiologia ambientale all'Imperial College di Londra e figura di spicco di "Human Genetics Foundation" di Torino, ora divenuta IIGM, *Italian Institute for Genomic Medicine*) le cui ricerche negli ultimi anni hanno permesso lo sviluppo di importanti innovazioni concettuali per la "genomica sociale". Vineis ha infatti originalmente contribuito allo sviluppo di alcuni fondamentali avanzamenti nelle conoscenze scientifiche della epidemiologia ambientale che gli hanno permesso di definire una delle ultime - omiche: l'esposoma (*exposomics*). Grazie all'impiego delle

tecniche più avanzate tra le *omiche* (metabolomica e epigenomica) ha studiato i biomarcatori capaci di segnalare i rischi di malattia dovuti all'esposizione a fattori ambientali nocivi, il che è di grande utilità per la medicina preventiva. Con il termine esposoma si intende infatti l'esposizione a quella miriade di fattori chimici e fisici (interni ed esterni) e psico-sociali (depressione, situazione economica, gerarchica, etc.) ai quali le cellule, i tessuti, gli organi ed i corpi sono costantemente soggetti in tutte le fasi dello sviluppo che nel loro insieme costituiscono la storia del ciclo vitale di un individuo (dall'embrione al senescente) e capaci di determinare stati metabolici (metabolomica) e cambiamenti funzionali a livello del genoma (epigenomica). La valutazione dell'esposoma (esterno: grazie all'uso di sensori e monitors personali; interno: grazie alle più moderne ed avanzate strumentazioni per omiche quali proteomica, metabolomica, epigenomica) è così in grado di fornire dei marcatori ambientali e fisiologici che possono essere trattati statisticamente per scoprire i meccanismi capaci di spiegare le condizioni di salute e di malattia e di mettere in luce le relazioni causali (o correlazioni) tra queste ed i contesti ambientali e socio-economici.

Questo approccio è stato premiato con prestigiosi riconoscimenti tra i quali spicca il finanziamento della Comunità Europea nell'ambito del programma Horizon 2020 per valutare il rischio di malattie dovute all'inquinamento di aria e acqua in diverse popolazioni europee. Di rilievo, anche per la impressionante mole di dati in uso, gli studi del mega-progetto Lifepath (stato socio-economico e invecchiamento in salute) coordinato dal Prof. Vineis e basati su metodologie sia osservazionali sia sperimentali. A partire dall'evidenza della vecchiaia in buona salute degli individui della più alta scala socio-economica rispetto a quelli posti nei ranghi inferiori, l'obiettivo finale del progetto è quello di fornire ai decisori politici i dati più aggiornati ed innovativi che permettano loro di elaborare politiche e strategie di inclusione sociale per colmare le disuguaglianze e promuovere una vecchiaia in buona salute per le più ampie fasce sociali possibili. Partendo dalle evidenze che già al concepimento, le condizioni fisiologiche sono in grado di influenzare lo stato di buona salute nel corso dell'invecchiamento, Vineis, in modo originale, teorizza e suddivide la storia del ciclo vitale dell'uomo in diverse fasi. La prima, è quella definita precoce (*early-life*, fase dello sviluppo, della costruzione), ove lo stato di salute determinato dai fattori ambientali e sociali a cui l'individuo è esposto peserà sull'invecchiare in buona salute; e una seconda fase, definita di declino in cui i fattori ambientali e sociali influenzano lo stato di salute della vecchiaia accelerando o meno il declino delle prestazioni funzionali, delle capacità fisiche e cognitive e lo sviluppo di patologie. Il divario in termini di mortalità tra i gruppi sociali è di tipo cumulativo e dunque le disuguaglianze aumentano con l'invecchiamento: i cambiamenti biologici dovuti al contesto sociale sono in buona parte reversibili (ed evitabili!) e dunque è dovere dei decisori politici fare buon uso delle tante fatiche dei ricercatori per fornire questo tipo di dati. Uno dei più recenti dei tanti lavori del gruppo guidato da Paolo Vineis è particolarmente interessante per i risultati acquisiti (Stringhini et al., 2015) e per la scrupolosa correttezza dello svolgimento delle ricerche (come è facile intuire, quello dell'epidemiologia sociale è un campo del sapere potenzialmente più fragi-

le di altri sotto il profilo dell'integrità dei ricercatori). Qui gli autori partendo dal fatto che un basso stato socioeconomico è associato ad una maggior produzione di cortisolo ed un aumentato stato infiammatorio si prefiggono di studiare la biologia molecolare di alcuni geni legati allo sviluppo dell'infiammosoma (capace di indebolire la risposta immunitaria a diversi patogeni). Studiando ben 857 individui sono stati in grado di dimostrare che il gene *NFATC1*, un gene che ricopre un ruolo chiave nella risposta immunitaria in presenza di patogeni (oltre a *IL1A* e *GPR132*, il malfunzionamento di questi due geni è legato allo sviluppo di infezioni e malattie autoimmuni) è significativamente "meno metilato"; al di là del gergo tecnico, è questa una modificazione epigenetica, funzionale, del gene e ciò significa che la sua espressione è molto ridotta negli individui con basso stato socioeconomico rispetto a quelli che occupano una posizione più alta nella scala sociale. Vineis dimostra così in modo inequivocabile che l'ambiente sociale è in grado di marcare in modo biologico (epigenetico) il genoma di un individuo: alcune di queste marcature epigenetiche sono poi capaci di passare attraverso un certo numero di generazioni future, di essere trasmesse alle generazioni future e così il vantaggio, più frequentemente lo svantaggio, ereditato passa alle nuove generazioni, le ingiustizie si perpetuano. Di rilievo il più recente "*The biological embedding of social differences in ageing trajectories*" (Vineis et al., 2016) ove è presentato in termini cristallini l'attuale drammatico scenario di malattie, perdita di funzionalità fisica e cognitiva nel corso dell'invecchiamento e mortalità per i gruppi svantaggiati sotto il profilo socioeconomico. Uno dei più recenti contributi di Paolo Vineis (2017) è un appassionato editoriale in difesa del principio di precauzione (ove contesta alcune delle posizioni che si stanno facendo strada tra coloro che ritengono che vi sia un'eccessiva cautela normativa che rallenta la produzione di beni e merci) che doverosamente permette anche di ricordare una figura prestigiosa di scienziato che ha speso la propria brevissima vita in difesa della salute pubblica, Giulio Alfredo Maccacaro (1924-1977). La rivista dell'associazione italiana di epidemiologia "Epidemiologia e Prevenzione" dedica nel suo primo numero dell'anno 2017 un ricordo speciale scritto da Enzo Ferrara (<http://www.epiprev.it/GAM/Intro>) alla figura di questo pionieristico scienziato della disciplina che oggi chiamiamo epidemiologia sociale. Alunno del Collegio Ghislieri di Pavia, si laurea in Medicina e Chirurgia nel 1948 e dopo diversi momenti di formazione in Europa (Cambridge *in primis*) fonda nel 1972 "Medicina Democratica", convinto assertore della prevenzione come fattore chiave per la tutela della salute dei lavoratori: aveva già intuito che il diritto alla salute nei luoghi di lavoro era un valore costitutivo della impresa, una premessa indispensabile della costituzione dei luoghi di produzione che deve precedere e non seguire l'organizzazione dei luoghi di lavoro. In altri termini, l'impresa deve organizzarsi sulla salute dei lavoratori; in fondo, esempi modello ne esistono a partire dalle concezioni di Adriano Olivetti per giungere alla attuale organizzazione produttiva della società Ferrero (sì, quella della mitica Nutella!): l'organizzazione del lavoro imprenditoriale deve essere basata sul principio di reinvestire a beneficio della comunità (dei lavoratori e della loro salute *in primis*) il profitto aziendale. Non è dunque un caso se nei giorni 9-10-11 novembre 2017 presso la Fondazione Fer-

ro di Alba si è tenuto un interessante convegno dedicato all'invecchiamento di successo e dove Paolo Vineis ha parlato di: "Determinanti sociali delle traiettorie di invecchiamento, epigenetica e *allostatic load*". Di rilievo, l'interesse per uno dei criteri più attuali e innovativi che la comunità scientifica utilizza in questi ultimi anni per misurare la capacità di un organismo di mantenere stabili le proprie funzioni fisiologiche in condizioni di stress: il carico allostatico. È questo il "prezzo" che il nostro organismo (cellule, tessuti, organi) "paga" per mantenersi in condizioni efficienti sotto condizione di mutamenti ed eventi di tipo stressante, anche in condizioni di cronicità. La grande innovazione in cui consiste la valutazione del carico allostatico è data dalla possibilità di rendere concreto e misurabile un concetto astratto come quello dello stress. In una precisa condizione di stress vengono valutati diversi parametri biochimico-clinici, tra i quali alcuni di quelli che si sono dimostrati più affidabili sono:

- pressione arteriosa sistolica;
- pressione arteriosa diastolica;
- emoglobina glicata;
- cortisolo urinario;
- colesterolo totale;
- epinefrina urinaria.

Questi ultimi vengono poi relazionati in una comparazione tra individui "stressati" e individui di riferimento, di controllo. La capacità di adattamento agli eventi spiacevoli della vita (dal lutto alla deprivazione del sonno, dall'isolamento sociale alla mancanza di cibo, etc.) è caratterizzata da una cascata di eventi che inizia con la risposta del sistema nervoso centrale ed è basata sia su fattori genetici sia su fattori esperienziali, in particolare dei primi anni di vita (deprivazioni emotive da parte dei genitori, abusi sessuali, etc.). Il carico allostatico è poi relazionato agli stili di vita ed è aumentato in modo significativo da scarso esercizio fisico, dieta ricca di grassi, fumo, abuso di alcool etc., tutti fattori significativamente tipici delle classi socioeconomiche più svantaggiate. Gli effetti negativi del carico allostatico vanno accumulandosi nel tempo e dunque saranno particolarmente dannosi nelle fasi dell'invecchiamento e della senescenza ove favoriscono in modo drammatico lo sviluppo di malattie.

L'ambiente inteso come sommatoria di fattori fisici, chimici, psicologici, sociali, etc. è in grado di determinare la nostra salute futura già marcando fasi molto iniziali dello sviluppo (embrionale, fetale ma anche cellule germinali; vedi oltre). In particolare, i primi nove mesi della nostra vita sono determinanti al fine del rimanente in buon stato di salute sino al momento in cui ciascuno di noi lascia il pianeta. In altri termini, le condizioni di sviluppo nel corso dei primi nove mesi di vita segnano in modo indelebile la nostra vita futura, ne modellano il rimanente biologico, tramite il contesto sociale vissuto dai nostri genitori. Lo svantaggio in salute della madre si rifletterà direttamente sulla salute del bimbo alla nascita e così lo svantaggio sociale diviene una ineguaglianza trasmessa alla generazione successiva. Un recente lavoro di Anna Aizer e di Janet Currie (2014) dichiara già dal titolo questa relazione: "*The intergenerational transmission of inequality*:"

Maternal disadvantage and health at birth". La salute del nascituro e del neonato è un importante predittore di salute sul lungo termine della storia del ciclo vitale di un individuo, non solo per aspetti attesi (disabilità) ma anche per aspetti meno ovvi quali la sua futura educazione/istruzione e la sua capacità di guadagno. Se il meccanismo di trasmissione è chiaro (condizione svantaggiata della madre - stili di vita non salutari - esposizione a fattori ambientali sfavorevoli - scarso accesso a cure mediche - peggiore salute del neonato) appare problematico il suo controllo e la sua eradicazione a causa del costante aumento, nel pianeta dei 7 miliardi di abitanti, di molte caratteristiche sociali di ingiustizia, dalla aspettativa di vita alla qualità della stessa.

Un corpo sempre più rilevante di evidenze scientifiche analizza in che modo l'espressione dei geni è influenzata dalle circostanze che definiscono la nostra vita quotidiana. Una sintesi di alto livello è proposta da Steven Cole (divisione di Ematologia-Oncologia, Università della California, Los Angeles) in un esaustivo lavoro di revisione della nuova disciplina "Genomica Sociale Umana" (Cole, 2014). Emerge dunque in modo chiaro come le condizioni socio-ambientali, tra le quali la collocazione urbana, il basso rango socio-economico, l'isolamento sociale e le minacce sociali per citarne alcune, correlino con una differenziale espressione di centinaia e centinaia di geni nei linfociti (i.e., cellule del sangue deputate tra le altre attività alle difese immunitarie) e nei tessuti patologici quali quelli invasi dalle metastasi di diverse forme tumorali. In tutte queste cellule tante e diverse condizioni di svantaggio socio-ambientale determinano un'aumentata espressione di geni pro-infiammatori ed una diminuita espressione di quelli deputati alle difese immunitarie. È stato così possibile mettere in evidenza i circuiti neuronali capaci di "trasdurre il segnale sociale in segnale genomico" legando specifici contesti sociali a precise condizioni funzionali del genoma (per una review vedi Cole, 2014). Queste evidenze hanno un doppio valore scientifico: da un lato sottolineano la relazione negativa tra specifici contesti socio-ambientali e salute e dall'altro empiricamente individuano quelle condizioni socio-ambientali più adatte al benessere ed allo sviluppo in salute degli individui suggerendo ai responsabili politici l'adozione di politiche sanitarie adatte ai singoli contesti sociali, storici, geografici, genetici e dello sviluppo che caratterizzano diversi gruppi di individui. La drammaticità di questa situazione non è solo sulla "carta" delle statistiche ma ben più tragicamente si rivela "sotto la pelle" delle grandi masse: basterà leggere l'ultimo rapporto sulle disuguaglianze nel mondo elaborato dall'organizzazione Oxfam (<http://www.oxfamitalia.org>). Inoltre, l'Organizzazione Mondiale della Sanità indica quale alta priorità delle nazioni "quella di assicurare un buon inizio della vita a ciascun bambino" (pag. 1012 di Marmot et al. "*WHO European review of social determinants of health and the health divide*" The Lancet 380:1011-1029, 2012) con la richiesta di ridurre entro poche generazioni le ineguaglianze di salute. I paesi, le istituzioni ed i singoli (fortunati) cittadini che non rispondano in modo appropriato a questo appello sono responsabili della degradazione e della umiliazione di persone che vedono cancellata, a causa del sociale che si fa biologico, la possibilità di realizzare il proprio progetto di sé (di perseguirlo e rivederlo nel corso del tempo) con le immagini spaventose ed inaccettabili di bimbi morti

con il volto nella sabbia di una spiaggia turca, morti perché costretti a migrare a causa delle tante guerre che l'Occidente sa creare. Purtroppo, proprio ora, dopo l'elezione di Donald Trump e la volontà di cancellare l'Obamacare, si sta negando ad almeno una ventina di milioni di cittadini statunitensi l'assistenza sanitaria senza la quale è ben difficile essere persone capaci di avere progetti ... e divenire miliardari in cinque mosse: la promessa elettorale di cancellare l'Obamacare per "costruire un meraviglioso (*"beautiful"*) sistema sanitario capace di servire ciascun singolo americano e che costerà meno assicurando di più" si sta rivelando per quello che era, una bugia, una classica *"fake-news"* che diviene vera per tanti americani perché ripetuta all'infinito dai media. Nell'era post-fattuale le *verità* si costruiscono e se non si hanno strumenti intellettuali e culturali per difendersi, beh, si vota Donald Trump e non Hillary Clinton.

Bibliografia essenziale

1. Aizer A, Currie J. The intergenerational transmission of inequality: Maternal disadvantage and health at birth, in "Science". 2014; 344: 856-861.
2. Cole S. Human Social Genomics, in "PloS Genetics". 2014, doi: 10.1371/journal.pgen.1004601.
3. Comitato etico fondazione Umberto Veronesi L'impatto delle disuguaglianze socio-economiche sul diritto ad avere eguali opportunità di salute in Italia. 2016. (https://www.fondazioneveronesi.it/uploads/2016/11/17/Comitato_Etico_Fondazione_Veronesi_Disuguaglianze.pdf)
4. Marmot M. La salute disuguale. La sfida di un mondo ingiusto, Il Pensiero Scientifico Editore, Roma. 2016.
5. McLean G, et al. The influence of socioeconomic deprivation on multimorbidity at different ages: a cross-sectional study, in "British Journal of General Practice". 2014; 64: e440-447.
6. Sherman G. et al. Leadership is associated with lower levels of stress, in "Proc. Natl. Acad. Sci. USA". 2012; 109: 17903-17907.
7. Stringhini S, et al. Life-course socioeconomic status and DNA methylation of genes regulating inflammation, in "Int. J. Epidemiology". 2015; 44: 1320-1330.
8. Underwood E. Can disparities be deadly?, in "Science". 2014; 344: 829-831.
9. Vineis P, et al. The biological embedding of social differences in ageing trajectories, in "J. Epidemiology Community Health". 2016; 70: 111-113.
10. Vineis P. Non rimettiamo in discussione il principio di precauzione, in "Epidemiologia e Prevenzione". 2017; 41: 6-7.

Riprogrammazione epigenetica delle cellule germinali

Manuela Monti

Laboratori di biotecnologie, Centro di Ricerche di Medicina Rigenerativa,
Fondazione IRCCS Policlinico San Matteo, Pavia

I gameti, oociti e spermatozoi, sono cellule altamente specializzate dalla cui fusione si genera la cellula staminale totipotente, ovvero lo zigote.

Tra i processi che portano alla formazione di tali cellule, la riprogrammazione epigenetica nelle cellule germinali primordiali (PGCs) e nelle successive fasi di sviluppo dell'embrione gioca un ruolo fondamentale per la loro specificazione.

Nel topo, le PGCs appaiono a circa 6.25 giorni dopo la fecondazione (E 6.25) e, grazie all'attivazione di fattori quali PRDM14 e AP2 γ , circa 40 PGCs migrano dalla base dell'allantoide alle creste genitali (E 11.5) e vanno incontro ad una importante ondata di riprogrammazione che comporta una globale demetilazione del DNA, cancellazione dell'imprinting genomico, riattivazione del cromosoma X e riorganizzazione delle modificazioni a carico della cromatina. Le cellule germinali maschili e femminili si differenziano al giorno E 13.5 e l'attivazione di diversi altri fattori quali BLIMP1 e STELLA ne determina la loro specificazione, insieme all'entrata in meiosi delle cellule germinali XX e all'arresto della mitosi nelle cellule germinali XY.

Nell'uomo, a differenza del topo, il processo che porta alla formazione di oogoni e spermatogoni (precursori di cellule uovo e spermatozoi maturi, rispettivamente) non è così ben conosciuto per evidenti problemi legati allo studio e dissezione di embrioni post-impianto (gastrule) ed è stato sostanzialmente estrapolato dal modello murino. Tuttavia, la comparsa della linea germinale è stabilita intorno alla seconda/terza settimana di sviluppo mentre durante la quarta settimana le PGCs localizzate nella parete del sacco vitellino, vicino all'allantoide, iniziano a migrare per colonizzare, verso la sesta settimana, le creste genitali dove proliferano sino alla decima settimana durante la quale le cellule germinali maschili entrano in quiescenza mitotica e quelle femminili si arrestano in profase meiotica.

Mentre le PGCs murine si specificano grazie alle proteine BMP2, BMP4 e WNT3 quando l'embrione post-impianto assume una struttura cilindrica caratterizzata dagli assi antero-posteriore e prossimale-distale, quelle umane (e di altri mammiferi non roditori) si presume si originino da una struttura planare in cui l'epiblasto più superficiale è caratterizzato dalla espressione di BMP4 e BMP8B e l'ipoblasto posteriore da BMP2 e WNT3 (Figura 1; Tang et al., 2016).

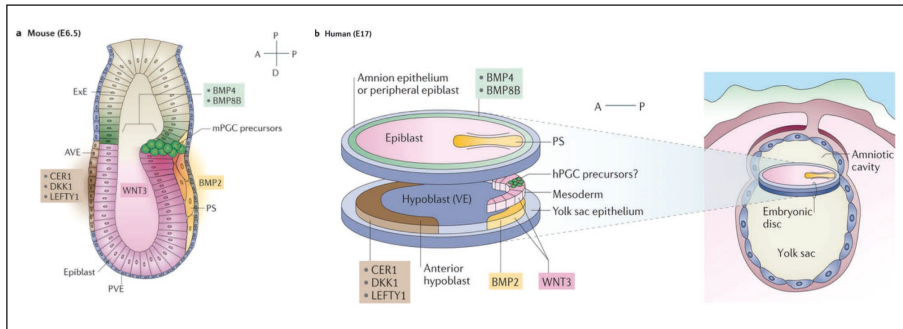


Fig. 1 - Formazione delle cellule germinali primordiali murine (a) e umane (b). ExE: ectoderma extra embrionale; AVE: endoderma viscerale anteriore; PVE: endoderma viscerale posteriore; PS: stria primitiva (da Tang et al., 2016).

Tralasciando i dettagli strettamente biologici del tanto complesso quanto affascinante processo della gastrulazione, definito dal biologo dello sviluppo Lewis Wolpert nel 1986 il momento più importante della nostra vita (Wolpert, 1991), ci focalizzeremo sulle recenti scoperte, in vitro ed in vivo, della specificazione delle PGCs umane e murine che sembrano avere una simile, sebbene non identica, dinamica di riprogrammazione epigenetica.

Nei mammiferi, durante lo sviluppo embrionale pre-impianto avviene una globale riprogrammazione genetica durante la quale l'epigenoma (l'insieme delle modificazioni della sequenza primaria del DNA) dello zigote viene resettato mentre la riprogrammazione nella linea germinale cancella la memoria epigenetica e facilita il processo di gametogenesi.

Nel topo, come detto in precedenza, le PGCs si originano da un epiblasto ipermetilato e per resettare l'epigenoma vanno incontro ad una estesa demetilazione del DNA (i livelli di metilazione delle isole CpG si riducono dal 70% a circa il 4%), riattivazione del cromosoma X, modificazione della cromatina per colonizzare le creste genitali, indicativamente dal giorno E7.5. Durante questa fase le regioni di controllo dell'imprinting sono ipometilate e la metilazione verrà di nuovo ristabilita in modo sesso-specifico al giorno E13.5 nei maschi e dopo la nascita nelle femmine. Si pensa che la globale demetilazione del DNA nella linea germinale si attui grazie ad un meccanismo passivo: poco dopo la specificazione delle PGCs, due fattori, PRDM1 e PRDM14 reprimono l'espressione delle DNA metiltransferasi 3A, 3B e della E3 ubiquitina protein ligasi UHRF1 (Magnusdottir et al., 2013) in modo che sia il mantenimento dei livelli di metilazione sia la nuova ondata di metilazione rimangano apparentemente repressi durante la proliferazione delle PGCs.

La conversione enzimatica di 5-metilcitosina (5mC) a 5-idrossimetilcitosina (5hmC) grazie alla famiglia degli enzimi TET (traslocazione dieci-undici, TET1, TET2, TET3), gioca un ruolo importante durante la demetilazione, in particolare nelle regioni imprimate, protette dalla demetilazione sino all'arrivo delle PGCs nelle creste genitali (Hackett et al., 2013). Sebbene studi su topi knockout abbiano mostrato che la demetilazione può verificarsi indipendentemente da TET 1 e

TET2, questi due fattori sono indispensabili per la cancellazione dell'imprinting e demetilazione dei promotori dei geni meiotici del differenziamento delle cellule germinali (Yamaguchi et al., 2012).

Nell'uomo, quando le PGCs colonizzano le creste genitali intorno all'ottava settimana, i livelli di metilazione si riducono al 4.5% e, come nel topo, la demetilazione è associata alla repressione di DNA metiltransferasi 3A, 3B e UHRF1. Tuttavia, a differenza del topo, sembra che l'onda di cancellazione dell'imprinting accada ben prima dell'arrivo delle PGCs nelle creste geniali (Tang et al., 2015).

La demetilazione del DNA nelle PGCs murine e umane è associata ad una globale riorganizzazione della cromatina. In particolare, la dimetilazione dell'istone H3 sulla lisina 9 (H3K9me2), una modificazione della cromatina associata a silenziamento, si riduce globalmente. Poiché vi è una sorta di interazione tra H3K9me2 e metilazione del DNA, la perdita di H3K9me2 nelle PGCs di topo può facilitare o essere una conseguenza della demetilazione del DNA. Un pattern simile è riscontrato anche nelle PGCs umane (Boroviak et al., 2014).

La metilazione del DNA viene di nuovo ristabilita durante la gametogenesi attraverso una metilazione *de novo* in oociti e spermatozoi insieme al rimodellamento della cromatina per garantire una corretta attivazione dell'oocita in previsione di un futuro sviluppo embrionale.

Vi sono profonde differenze nell'organizzazione epigenetica di oociti maturi e spermatozoi il cui DNA è molto più metilato di quello degli oociti. Alla fecondazione, i genomi di oociti e spermatozoi sono trascrizionalmente silenti mentre negli istanti seguenti i cromosomi degli spermatozoi si decondensano e rimodellano, le protammine si sostituiscono agli istoni e avviene una rapida demetilazione. Il genoma degli oociti è demetilato da un meccanismo passivo ma i geni imprintati rimangono protetti dalla nuova ondata di demetilazione in modo da conservare l'imprinting parentale. Durante l'impianto avviene una nuova ondata di metilazione e la riprogrammazione epigenetica diviene essenziale per un corretto sviluppo regolando l'espressione genica durante le prime fasi di sviluppo embrionale, divisione e specificazione cellulare. Durante le fasi più tardive di sviluppo embrionale, i geni embrionali vengono silenziati a favore dell'espressione di geni tessuto-specifici. Una incompleta o aberrante riprogrammazione epigenetica dell'embrione preimpianto può comportare ritardi nello sviluppo embrionale o, in casi più gravi, a morte. Diversamente da modificazioni genetiche, alcune alterazioni epigenetiche possono essere tollerate durante lo sviluppo embrionale e generalmente sono reversibili come risultato della cancellazione nella linea germinale (Jacob e Moley, 2005).

Ambiente e nutrizione: qualità delle cellule germinali

Vi è un aspetto assai importante che deve necessariamente essere considerato quando si parla di epigenetica, gameti e sviluppo embrionale e che si riferisce alla informazione che viene passata dai genitori alla progenie e che, al di là dei fattori più strettamente genetici (alcuni dei quali sono stati illustrati sopra), riguarda l'influenza dell'ambiente sull'imprinting genetico. Quest'ultimo condiziona per

via epigenetica gli individui in tutti i momenti della loro storia del ciclo vitale. Quando sono i tessuti somatici di un adulto (a qualunque livello di sviluppo, da embrione a senescente) ad essere epigeneticamente condizionati, la nuova costituzione genetica si esprime in quella sola generazione. Quando però l'ambiente è in grado di marcare per via epigenetica (metilazione del DNA, acetilazione delle proteine, etc.) il genoma delle cellule germinali, allora si realizza la possibilità che quelle modificazioni siano passate alla generazione (o a più generazioni) successiva, nel momento in cui l'individuo portatore di quelle cellule germinali si riproduce. Come facilmente intuibile un aspetto di estrema rilevanza ai fini di una corretta riproduzione è quello che lega la qualità delle cellule uovo (in termini di corretto sviluppo embrionale e quindi fetale, e quindi dello sviluppo del neonato) allo stato di nutrizione della femmina che lo produce. Se sterminata è la bibliografia su quello che già intuitivamente è un aspetto deleterio e negativo per la qualità dell'ocita, la sotto-nutrizione (tragica e dolorosa realtà che riguarda decine e decine di milioni di donne), meno conosciuto e meno intuitivamente pericoloso è quello del legame tra obesità, sovrappeso in ogni caso, e qualità dell'ocita. La sovra-nutrizione nelle femmine causa un alterato sviluppo fetale della progenie e programma stabilmente in modo negativo il metabolismo dei nati. E' stato infatti dimostrato nei topolini che gli oociti prodotti da femmine obese mostrano caratteristiche molecolari fuori dal normale assetto; ad esempio il citoplasma di questi oociti è infarcito di goccioline lipidiche, il reticolo endoplasmatico si presenta ipertrofico, le fibre del fuso delle anomalie di struttura, e sebbene questi oociti contengano un normale numero di mitocondri (organelli presenti nel citoplasma e dedicati alla produzione di energia sotto forma di legami chimici ad alto valore energetico), le caratteristiche fisiologiche di questi ultimi non sono regolari e presentano un alto tasso di autodistruzione (autofagia) rispetto agli oociti prodotti da femmine normo-peso. E così, dopo fecondazione in vitro, questi oociti mostrano uno scarso potenziale di sviluppo embrionale arrivando a formare delle blastocisti (uno delle fasi precoci dello sviluppo preimpianto) con ridotti numeri di mitocondri (Wu et al., 2015). Una volta che gli embrioni derivanti dalle blastocisti prodotte da oociti di madri obese sono trasferiti in pseudo madri, sviluppano feti di maggior peso corporeo rispetto ai controlli formati da madre di regolare peso. Ciò che è grave è il fatto che questi feti hanno cellule del fegato e del rene con un minore contenuto di mitocondri, con ciò indicando che l'obesità materna, prima della fecondazione, altera la trasmissione dei mitocondri alla progenie: è bene ricordare che nei Mammiferi la trasmissione di questi particolari organelli citoplasmatici è solo materna. Oggi conosciamo delle molecole capaci di recuperare questa situazione: il trattamento delle madri obese con Salubrial (un inibitore della deregolazione della corretta composizione/organizzazione del citoplasma degli oociti) riporta a livelli normali il contenuto di mitocondri e corregge le malformazioni della struttura del citoplasma. Il trattamento farmacologico, con la sua efficacia, funziona comunque da controprova della negativa influenza esercitata dalla sovra-nutrizione sul processo di oogenesi (Wu et al., 2015). Tutti questi dati sono stati ottenuti grazie a sperimentazione scientifica su modelli murini: sostengono del tutto i dati epidemiologici di ciò che analogamente accade nelle donne; quelle

in sovrappeso ricorrono ben più frequentemente alle tecniche di fecondazione assistita, ad esempio alla fecondazione in vitro, a causa di alterati stati di fertilità (sub-fertilità) e comunque il successo delle tecniche di fecondazione in vitro è minore rispetto a quello che si registra nelle pazienti di normale peso corporeo, a causa di uno scarso potenziale di sviluppo embrionale degli oociti. Il legame tra le disfunzioni mitocondriali delle cellule uovo indotte dall'eccessiva assunzione di cibo, dall'obesità, e lo sviluppo della resistenza all'insulina (diabete) ed allo sviluppo dell'obesità nel corso dello sviluppo embrionale e fetale è stato recentemente oggetto di una revisione sistematica della letteratura giungendo alla conclusione che non solo nei modelli animali ma anche nella nostra specie madri obese (e comunque in sovrappeso) predispongono i nascituri a diabete e obesità (Leary et al., 2015; Turner e Robker, 2015). L'evoluzione delle abitudini alimentari verso il consumo di cibi estremamente lavorati, ricchi di zucchero e acidi grassi saturi, che ha portato all'attuale pandemia di sovrappeso e obesità, ha trascinato con sé anche uno smodato consumo di cibi ad alto contenuto salino. È così emerso che anche l'eccesso di sale è dannoso per lo sviluppo degli oociti: un recente lavoro ha messo in evidenza come una dieta particolarmente ricca di sale influisca negativamente sulla fertilità femminile (Wang et al., 2015). In particolare, è stato dimostrato che in condizioni di regime alimentare ipersalino la popolazione degli oociti primordiali (quelli di dimensioni più piccole, ovvero all'inizio del processo maturativo dell'oogenesi) viene danneggiata poiché si altera il rapporto nutritivo tra cellula uovo e le cellule della granulosa che ne accompagnano lo sviluppo sino ad oocita antrale pronto per la ovulazione. In altri termini gli oociti primordiali non sono in grado di progredire nella crescita sino a raggiungere lo stadio di oociti primari e, al contrario, arrestano la loro maturazione. Tra le molte cause che portano a questo arresto, una è data dalla atresia e dalla inibizione della proliferazione cellulare delle cellule somatiche che circondando l'oocita, quelle che costituiscono il follicolo, ovvero quella nicchia indispensabile per la corretta crescita dell'oocita costituendo il veicolo di scambi di ormoni, fattori di crescita e sostanze dall'ambiente esterno a quello interno.

Se è ben noto quanto l'iperglicemia materna e la obesità durante la gestazione siano fattori di rischio per la capacità di indurre diabete e obesità nella progenie, recenti lavori sostengono che anche il profilo metabolico del padre al momento del concepimento sia in grado di influire sul profilo metabolico della progenie: uomini obesi concepiscono bambini obesi (Li et al., 2009). Più fattori associati all'obesità sono in grado di compromettere il processo di spermatogenesi ed in particolare l'accumulo di grasso a livello dello scroto e della zona sovrapubica poiché questa condizione anatomica determina un aumento della temperatura del testicolo e ciò danneggia in modo irreversibile sia la produzione (con la comparsa di tubuli seminiferi atrofici) sia il funzionamento degli spermatozoi che vedono alterarsi parametri strutturali quali la conformazione della cromatina sia parametri funzionali quali il processo di capacitazione degli spermatozoi attraverso il quale si realizza l'acquisizione della capacità di legarsi alla cellula uovo. Anche la motilità degli spermatozoi è ridotta drasticamente mentre aumenta in modo statisticamente significativo il numero degli spermatozoi aberranti (terato-

spermatozoi); anche l'integrità della barriera emato-testicolare è compromessa. È evidente da questa prima indicazione che sia necessario mettere in campo azioni di informazione e sorveglianza prescrivendo condizioni di normo-peso anche ai maschi, perlomeno nel periodo in cui siano impegnati in relazioni di paternità. In uno studio ben condotto ove sono state analizzate ben 651 coppie in regime di fecondazione assistita, è risultato ben chiaro che la più bassa capacità fecondante è quella che caratterizza i maschi con un indice di massa corporea superiore a 28 kg/m²; la più bassa capacità fecondante comporta inevitabilmente una più bassa qualità degli embrioni derivati e utili all'impianto in utero (Yang et al., 2016). Questo quadro clinico è dovuto al fatto che un maggior numero di spermatozoi di maschi in sovrappeso ed obesi presentano diverse caratteristiche negative: - frammentazione del DNA;

- presenza di radicali liberi di ossigeno (a causa di stress ossidativo) ad alta reattività chimica;
- telomeri ridotti in lunghezza;
- mitocondri danneggiati.

Le relazioni tra obesità, bilancia energetica, spermatogenesi ed in ultima analisi condizione di fertilità maschile (sterilità, subfertilità) sono oggi ben note: una dieta altamente calorica porta facilmente ad un aumento del peso corporeo e ad iperglicemia, iperinsulinemia e dislipidemia. Questo stato fisiologico causa una infiammazione cronica ed un alto metabolismo di base (necessario a sostenere la bilancia metabolica). L'alto metabolismo di base così sviluppato innalza i livelli di molecole di ossigeno reattivo (radicali liberi di ossigeno, i famosi ROS, "reactive oxygen species") capaci di danneggiare il DNA, le proteine e l'integrità delle membrane cellulari. Questi eventi molecolari e citologici risultano gravi per tutti i tipi di tessuto e di organi, soprattutto per il testicolo per la particolare sensibilità delle cellule germinali alla presenza di ROS (Jesus et al., 2016).

Per concludere dunque, non solo la nutrizione ma anche lo stile di vita della madre può influenzare lo sviluppo embrionale portando a modificazioni epigenetiche sul feto: condizioni di stress, esposizione a sostanze chimiche, fumo, consumo di alcool e droghe possono causare modificazioni geniche destinate al controllo metabolico o endocrino dell'organismo aumentando la predisposizione del feto a malattie nell'età adulta quali malattie cardiovascolari, sindrome metaboliche e diabete, solo per citarne alcune (Boekelheide et al., 2012).

Bibliografia

1. Boekelheide K, Blumberg B, Chapin RE, Cote I, Graziano JH, Janesick A, Lane R, Lillycrop K, Myatt L, States C, Thayer KA, Waalkes MP, Rogers JM. Predicting Later-Life Outcomes of Early-Life Exposures. *Environ Health Perspect.* 2012; 120: 1353-1361.
2. Boroviak T, Loos R, Bertone P, Smith A, Nichols J. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat Cell Biol.* 2014; 16: 516-528.

3. Hackett JA, Sengupta R, Zyllicz JJ, Murakami K, Lee C, Down T, Surani MA. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science*. 2013; 339: 448-452.
4. Jacob S, Moley KH. Gametes and Embryo Epigenetic Reprogramming Affect Developmental Outcome: Implication for Assisted Reproductive Technologies. *Pediatric Res*. 2005; 58: 437-446.
5. Jesus TT, Oliveira PJ, Silva J, Barros A, Ferreira R, Sousa M, Cheng Y, Silva BM, Alves MG. Mammalian target of rapamycin controls glucose consumption and redox balance in human Sertoli cells. *Fertil Steril*. 2016; 105: 825-833.
6. Leary C, Leese HJ, Sturmey RG. Human embryos from overweight and obese women display phenotypic and metabolic abnormalities. *Hum Reprod*. 2015; 30: 122-132.
7. Li L, Law C, Lo Conte R, Power C. Intergenerational influences on childhood body mass index: the effect of parental body mass index trajectories. *Am J Clin Nutr*. 2009; 89: 551-557.
8. Magnusdottir E, Dietmann S, Murakami K, Günesdogan U, Tang F, Bao S, Diamanti E, Lao K, Gottgens B, Surani MA. A tripartite transcription factor network regulates primordial germ cell specification in mice. *Nat Cell Biol*. 2013; 15: 905-915.
9. Tang WW, Dietmann S, Irie N, Leitch HG, Floros VI, Bradshaw CR, Hackett JA, Chinnery PF, Surani MA. A unique gene regulatory network resets the human germline epigenome for development. *Cell*. 2015; 161: 1453-1467.
10. Tang WW, Kobayashi T, Irie N, Dietmann S, Surani MA. Specification and epigenetic programming of the human germ line. *Nat Rev Genet*. 2016; 17: 585-600.
11. Turner N, Robker RL. Developmental programming of obesity and insulin resistance: does mitochondrial dysfunction in oocytes play a role? *Mol Hum Reprod*. 2015; 21: 23-30.
12. Wang G, Yeung CK, Zhang JL, Hu XW, Ye YX, Yang YX, Li JC, Lee KK, Yang X, Wang LJ. High salt intake negatively impacts ovarian follicle development. *Ann Anat*. 2015; 200: 79-87.
13. Wolpert L. *The triumph of the embryo*. Oxford University press. 1991; 221.
14. Wu LL, Russell DL, Wong SL, Chen M, Tsai TS, St John J, Norman RJ, Febbraio MA, Carroll J, Robker RL. Mitochondrial dysfunction in oocytes of obese mothers: transmission to offspring and reversal by pharmacological endoplasmic reticulum stress inhibitors. *Development*. 2015. 142: 681-691.
15. Yamaguchi S, Hong K, Liu R, Shen L, Inoue A, Diep D, Zhang K, Zhang Y. Tet1 controls meiosis by regulating meiotic gene expression. *Nature*. 2012; 492: 443-447.
16. Yang Q, Zhao F, Hu L, Bai R, Zhang N, Yao G, Sun Y. Effect of paternal overweight or obesity on IVF treatment outcomes and the possible mechanisms involved. *Sci Rep*. 2016; 6: 29787.

Pathways to neurodegeneration: beyond the genome

Enza Maria Valente

Department of Molecular Medicine, University of Pavia

The complex pathogenesis of Parkinson disease: genetics and beyond

Parkinson Disease (PD) is one of the commonest neurodegenerative disorders, with prevalence over 1% in the seventh decade of life. The phenotype of bradykinesia, resting tremor, rigidity and postural instability is mainly caused by the massive death of dopaminergic neurons in the substantia nigra pars compacta. Surviving neurons show the Lewy bodies (LBs), typical cytoplasmic inclusions containing ubiquitin, alpha-synuclein (α -syn) and other proteins. Despite familial aggregation had long been recognized as a common feature in PD, only in the past fifteen years the contribution of genetics has been deeply explored, with the identification of few genes clearly responsible for mendelian forms of the disease, either with autosomal dominant (e.g. SNCA, LRRK2) or recessive (e.g. PARK2/Parkin, PINK1, DJ-1, ATP13A2) inheritance. Besides these, some other genes have been found mutated only in rare families, and their actual contribution to the disease remains to be confirmed.

However, it was recently estimated that mutations in known genes can only explain about 10% of familial PD, leaving a large proportion of cases unaccounted. It is conceivable that in the vast majority of these patients, as well as in those with negative family history, the disease underlies a multifactorial inheritance, with several common and rare genetic variants interplaying among them and with environmental factors to reach a threshold of disease. In fact, heterozygous mutations in the GBA gene were found to represent a strong genetic risk factor for development of PD, with an average relative risk of about 5, and several polymorphisms (mostly residing within the same genes mutated in mendelian forms of PD) were identified as PD susceptibility factors by large whole genome association (GWA) studies. In line with this hypothesis, the advent of next generation sequencing (NGS) technologies, which have impressively enhanced the identification of novel causative genes in most genetic disorders, has not brought the expected revolution in the field of PD genetics. Indeed, despite NGS has been available for over 6 years, only few rarely mutated genes have been found in isolated families, and a single study has been published reporting whole exome sequencing in 100 Sardinian patients with PD, which failed to identify robust novel PD candidates.

While it can be safely expected that the huge amount of genetic data generated through NGS and GWA strategies contains a wealth of new relevant information

to better understand the complex genetic basis of PD, nevertheless it is undoubted that the meaningful analysis of such data to extrapolate novel PD candidate genes represents a truly challenging task. A possible lead to get oriented in the maze of common and rare genetic variants emerging from genome-wide studies may derive from the analysis of pathogenic mechanisms of neurodegeneration. It is clearly emerging that neuronal death results as a consequence of the derangement of very few essential and highly interconnected cellular pathways, and that many if not all genetic factors implicated in PD pathogenesis converge on, and affect, the same cellular processes, including mitochondrial dysfunction, misfolded protein damage, impairment of clearance systems, abnormal calcium handling and enhanced pro-inflammatory responses. Intriguingly, mounting evidence is now showing that epigenetic mechanisms affecting the expression levels of key PD genes play a central role in the triggering of the neurodegenerative process, contributing to explain the individual susceptibility to PD and variability of phenotypes, as well as the variable penetrance of most PD-related genetic factors. Here we will discuss this complex interplay between genetic and epigenetic factors in PD pathogenesis by taking the most important and relevant example: the one of α -syn.

Focus on SNCA/ α -syn

The central role of α -syn in PD pathogenesis emerged twenty years ago, with the identification of the p.A53T mutation as the first genetic cause of inherited, autosomal dominant PD in an Italian-Greek family, the Contursi kindred. A year later, this previously neglected protein was found in its aggregated form to be the most abundant component of Lewy bodies, Lewy neurites and glial cell inclusions, thereby creating a direct and unexpected link between PD, multiple system atrophy and Lewy Body Dementia (LBD), which were then collectively termed “synucleinopathies”.

The same p.A53T mutation was subsequently found in several Greek/Italian families all sharing a common ancestor, in two other PD families of Korean and Swedish origin and in an apparently sporadic Polish case. The phenotype of mutated patients carrying the p.A53T mutation ranged from typical late-onset PD to atypical PD with more severe features, such as earlier age at onset, rapid progression, and high prevalence of cognitive, psychiatric and autonomic impairment. Other few missense *SNCA* mutations have been reported in few familial cases with PD (with relative early onset and rather aggressive course) or LBD. At difference from point mutations, that are extremely rare, *SNCA* locus multiplications represent a more frequent cause of autosomal dominant PD. Triplications of the whole gene have been described in few families presenting a severe form of early onset parkinsonism, while *SNCA* duplications are even commoner, being reported in several familial and sporadic PD cases worldwide. In this genetic condition, the disease severity appears to correlate well with the dosage of the *SNCA* gene, rather than with the extension of the multiplied genomic region. Indeed, the presence of four *SNCA* copies is always causative of a fully penetrant, aggressive

and rapidly progressive phenotype, with early onset (usually in the third to fourth decade) of parkinsonian signs, and precocious non-motor features (dementia, psychiatric disturbances and dysautonomia). Conversely, in patients with SNCA duplications the disease presentation is highly variable, even within families: in some patients it resembles idiopathic late-onset PD, while others are more similar to patients bearing SNCA triplications. Penetrance is estimated to be up to 30%, as several healthy carriers of the SNCA duplication have been reported.

α -syn is a 140 amino-acid protein encoded by the SNCA gene on chromosome 4q22.1. It is highly expressed in neural tissues, representing more than 1% of the total brain protein content, where it predominantly localizes to presynaptic terminals. Although its physiologic function is not fully characterized yet, α -syn has been implicated in synaptic plasticity, vesicle dynamics and dopamine metabolism. Moreover, a large wealth of functional studies on *in vitro* and *in vivo* models have unequivocally demonstrated that accumulation and aggregation of wild type α -syn has a deleterious impact on a number of key mechanisms, including mitochondrial homeostasis and dynamics, endoplasmic reticulum stress and functioning of the two major cellular clearance pathways, namely the ubiquitin-proteasome system and autophagy and, as expected, all identified SNCA pathogenic mutations were found to increase α -syn toxicity.

Epigenetic regulation of the SNCA gene

Recent efforts have shed new light on the epigenetic mechanisms driving gene expression alterations associated with PD pathogenesis. Changes in gene expression are a well-established cause of PD, and epigenetic mechanisms likely play a pivotal role in regulation. Of note, genome-wide association studies have identified over 30 loci associated with PD with most signals located in non-coding regions of the genome, thus likely influencing transcript expression levels.

Although the role of DNA methylation and its link to PD pathogenesis is currently unclear, methylation of the SNCA gene may result in protein overexpression or altered gene expression. Methylation of intron 1 was associated with decreased SNCA transcription, whereas reduced methylation was found to be decreased in several brain regions, including the substantia nigra, of sporadic PD patients, causing increased SNCA expression. Moreover, α -syn can sequester DNA methyltransferase 1 in the cytoplasm (needed to maintain DNA methylation), leading to global DNA hypomethylation in PD and LBD, as well as in α -syn mouse models [25]. Conversely, overexpression of DNMT1 could restore normal protein levels, underscoring the significance of epigenetic dysregulation in PD.

SNCA expression can also be altered by histone modifications. In *Drosophila* models, nuclear-targeted α -syn was shown to bind to histones and reduce histone H3 acetylation through its association with the HDAC SIRT2. Consistent with these findings, the targeted downregulation of SIRT2 could rescue α -syn toxicity and dopaminergic cell loss in flies and primary mesencephalic cultures. Moreover, α -syn toxicity in both cellular and animal models could be rescued by HDAC inhibitors.

Finally, several miRNAs and their corresponding targets have been identified as regulators of SNCA function, providing new mechanistic insights regarding the possibility of using RNA interference (RNAi) as a therapeutic approach. For instance, miR-7 and miR-153 were found to possess target sites in the 3'UTR of SNCA. miR-7 is mainly expressed in neurons and can repress α -syn protein levels via the 3'UTR of SNCA mRNA. Mutant α -syn-induced toxicity could be attenuated by miR-7, suggesting that miR-7 plays a neuroprotective role. miR-153 is another miRNA known to repress α -syn expression at both the mRNA and protein levels. Although the functional significance of these altered miRNA levels have yet to be established, miRNAs could be explored as therapeutic targets to block α -syn-induced abnormalities in PD.

Conclusions

Over recent years, there has been robust evidence generated showing that epigenetic dysregulation occurs in PD patients, and that epigenetic modulation is a promising therapeutic approach for PD. Moreover, the epigenome holds considerable promise for the development of molecular biomarkers for PD. Epigenetic marks are modified by both DNA sequence and environmental factors associated with PD, and such marks could serve as a unifying predictor of at-risk individuals, expanding current opportunities for personalized therapeutic strategies.

Bibliografia essenziale

1. Deng H, Yuan L. Genetic variants and animal models in SNCA and Parkinson disease. *Ageing Res Rev.* 2014; 15: 161-76.
2. De Rosa P, Marini ES, Gelmetti V, Valente EM. Candidate genes for Parkinson disease: Lessons from pathogenesis. *Clin Chim Acta.* 2015; 449: 68-76.
3. Guhathakurta S, Bok E, Evangelista BA, Kim YS. Deregulation of α -synuclein in Parkinson's disease: Insight from epigenetic structure and transcriptional regulation of SNCA. *Prog Neurobiol.* 2017; 154: 21-36.
4. Labbé C, Lorenzo-Betancor O, Ross OA. Epigenetic regulation in Parkinson's disease. *Acta Neuropathol.* 2016; 132 :515-30.
5. Petrucci S, Consoli F, Valente EM. Parkinson Disease Genetics: A "Continuum" From Mendelian to Multifactorial Inheritance. *Curr Mol Med.* 2014; 8: 1079-88.
6. Petrucci S, Ginevrino M, Valente EM. Phenotypic spectrum of alpha-synuclein mutations: New insights from patients and cellular models. *Parkinsonism Relat Disord.* 2015; 22(Suppl. 1): S16-20.

Epigenomica e malattie neurodegenerative

Cristina Cereda

Genomic and post-Genomic Center, IRCCS “C. Mondino”
National Institute of Neurology Foundation, Pavia

Nel 1940 Conrad Waddington coniò il termine “epigenesis” per descrivere l’insieme di meccanismi biologici che si collocano sopra ‘epi’ il livello del DNA e che regolano l’espressione differenziale di geni nei diversi tipi cellulari determinandone il destino (1). L’epigenetica comprende quindi un’ampia gamma di cambiamenti ereditabili che non risultano da alterazioni nella sequenza del DNA, ma che determinano variazioni nell’espressione di geni in risposta all’ambiente. I meccanismi epigenetici di regolazione dell’espressione genica si possono verificare mediante il “rimodellamento della cromatina”, attraverso la metilazione o idrossimetilazione del DNA, oppure mediante modifiche post-traduzionali a carico degli istoni con conseguente riposizionamento del nucleosoma. Altri meccanismi epigenetici sono mediati da microRNA (miRNA) e da RNA lunghi non codificanti (lncRNA) (2-7). I meccanismi epigenetici regolano l’espressione genica, il differenziamento cellulare e lo sviluppo di quasi tutti i tessuti, incluso il cervello. Dallo sviluppo embrionale all’età adulta i cambiamenti nell’epigenoma sono cruciali per le funzioni cognitive superiori come l’apprendimento e la memoria e nuove evidenze ne sottolineano l’implicazione nell’insorgenza delle malattie neurodegenerative.

La *metilazione del DNA* prevede l’aggiunta covalente di un gruppo metilico dall’S-adenosil metionina (SAM) al carbonio 5 della citosina grazie all’azione degli enzimi DNA metiltrasferasi (DNMTs).

In questo modo si forma la 5-metilcitosina (5-mC) con la concomitante conversione di SAM in S-adenosilomocisteina (SAH) (8). La metilazione è un processo dinamico che coinvolge vari siti genomici anche se è stato descritto principalmente nelle ripetizioni dei dinucleotidi CG raggruppati in aree note come “isole CpG” localizzate nelle regioni promotrici dei geni determinando l’inibizione della trascrizione (9, 10).

L’*idrossimetilazione* è una modifica del DNA dovuta all’ossidazione di 5-mC a 5-idrossimetilcitosina (5-hmC) mediante l’azione della proteina TET1 (11). La 5-hmC è spesso associata a regioni euromatiche, indicando la sua relazione con l’aumento della trascrizione (12, 13).

Le *modifiche istoniche* coinvolgono le code N-terminali degli istoni lunghe circa 25-40 residui amminoacidici, dove gli enzimi che modificano la cromatina possono eseguire la loro funzione. Le modifiche degli istoni includono la metilazione dei residui di lisina o arginina, l’acetilazione, la fosforilazione, l’ubiquitina-

zione, la SUMOilazione, l'ADP-ribosilazione, la crotonilazione, l'idrossilazione e l'isomerizzazione della prolina (14).

Anche gli *RNA non codificanti (ncRNA)*, in particolare gli RNA lunghi non codificanti (lncRNA) e i microRNA (miRNA) sembrano svolgere un ruolo regolatore fondamentale in molti processi cellulari fisiologici e patologici (ad esempio nei tumori e nella neurodegenerazione) (15). I miRNA si legano alla regione 3' UTR dell'mRNA dei geni bersaglio e ne modulano la trascrizione (16), mentre per i lncRNA, una classe di ncRNA con più di 200 nucleotidi (17, 18), il contributo al processo di regolazione è molto più complesso e ancora in gran parte sconosciuto. È noto oramai come i lncRNA siano spesso sovra-espressi nel cervello e coinvolti in molti disturbi neurodegenerativi (19, 20). I lncRNA sono tessuto specifici e tendono a essere espressi a livelli simili a quelli degli RNA codificanti corrispondenti; non è ancora nota la loro origine ma si pensa che questi derivino da inserzioni e delezioni di elementi trasponibili (21). A oggi sono stati identificati più di 13500 lncRNAs umani, suddivisi in quattro categorie: long intergenic non-coding RNAs (lincRNAs), sense lncRNAs, trascritti nello stesso senso del gene codificante, lncRNAs bidirezionali o divergenti, e lncRNAs antisense (AS-lncRNA), trascritti nella direzione opposta rispetto al gene codificante. Tutti questi lncRNA sono coinvolti nei processi di regolazione post-trascrizionale (22, 23) e in condizioni fisiologiche sono stati associati a diverse funzioni tra cui lo sviluppo e la differenziazione neuronale (24, 25), la plasticità sinaptica, le funzioni cognitive e i processi mnemonici.

Epigenetica nella malattia di Parkinson

La malattia di Parkinson (PD) è un disturbo complesso causato dall'interazione tra fattori genetici, ambientali, nutrizionali e legati all'invecchiamento. Poiché l'epigenetica può essere alterata in risposta ad alcuni di questi fattori, sta diventando sempre più evidente che le variazioni epigenetiche possano svolgere un ruolo importante sia nell'eziologia sia nella patogenesi della malattia.

L'*analisi di metilazione del DNA* effettuata su tutto il genoma in campioni di sangue e di cervello di pazienti affetti da PD ha evidenziato un incremento significativo della metilazione nel gruppo di pazienti rispetto a individui sani. Molti geni sono stati individuati come ipo- o iper-metilati, compresi i geni responsabili della suscettibilità alla malattia (26, 27). Il promotore del gene *SNCA* è solitamente ipometilato nel PD, con conseguente aumento dei livelli di α -sinucleina (α Syn). Inoltre l' α Syn è in grado di sequestrare nel citoplasma DNMT1 causando una riduzione generale del pattern di metilazione (28). Nei campioni cerebrali *post-mortem* di PD è stata riscontrata una metilazione aberrante in alcuni geni come *PARK16*, *GPNMB* e *STX1B* e una diminuzione della metilazione con conseguente aumento dei livelli di mRNA del gene *CYP2E* (29). Sebbene i promotori di *PARK2*, *UCHL1* e *ATP13A*, geni mutati nella malattia di Parkinson, presentino iper-metilazione in alcuni tumori, nella malattia di Parkinson non sono stati descritti livelli anormali di metilazione (30-32). Le *modifiche degli istoni* svolgono un ruolo chiave nello sviluppo, differenziazione e mantenimento dei neuroni dopaminergici (33), ma

poco si sa riguardo il loro coinvolgimento nella patogenesi del PD. In uno studio recente, l'uso di neuroni dopaminergici isolati dal tessuto cerebrale di pazienti PD ha rilevato un aumento dei livelli di acetilazione degli istoni H2A, H3 e H4 rispetto ai soggetti di controllo di pari età. Inoltre, i livelli degli enzimi responsabili della deacetilazione, le istone deacetilasi (HDACs), sono ridotti nel cervello di topi trattati con MPTP (1-metil 4-fenil 1,2,3,6-tetraidro-piridina), una tossina utilizzata per indurre sperimentalmente il PD (34). I neuroni dopaminergici dei topi trattati con paraquat, un altro agente neurotossico che induce acetilazione dell'istone H3, presentano un accumulo di α Syn nel nucleo, dove si co-localizza con l'istone H3 acetilato (35). Ulteriori indagini hanno rivelato che l' α Syn si lega direttamente all'istone H1 formando un complesso stabile. Inoltre, l'istone H1, insieme agli istoni principali, è in grado di aumentare la formazione di fibrille di α Syn (36). Anche i *miRNA* sono stati oggetto di attenzione nello studio del PD. Il miR-133b, specifico per i neuroni dopaminergici del mesencefalo sembra essere down-regolato nei pazienti PD. Il miR-132 anch'esso collegato al differenziamento neuronale dopaminergico del mesencefalo è invece significativamente aumentato in modelli animali (ratti) di PD e i livelli della sua proteina bersaglio, Nurr1, sono ridotti (37, 38). Studi di espressione hanno trovato livelli di miRNA differenzialmente espressi tra pazienti e controlli: i livelli di miR-1, miR-22 e miR-29 sono ridotti in campioni di sangue di pazienti PD, mentre il miR-1826, miR-450b-3p, miR-505 e miR-626 sono sovra-regolati e possono essere utilizzati come biomarcatori della patologia (39). Inoltre, è stato trovato che il miR-205 è in grado di legarsi al 3'UTR dell'mRNA di LRRK2 aumentando l'espressione; i campioni di corteccia frontale di pazienti PD contengono elevati livelli di LRRK2 e livelli ridotti di miR-205. Infine, l'introduzione del miR-205 in neuroni portatori della mutazione R1441G *LRRK2* impedisce l'insorgenza di difetti di crescita (40). I miR-7 e miR-153 si legano al 3'UTR dell'mRNA del gene *SNCA* inibendone la traduzione. In particolare, il miR-7, specifico dei neuroni, diminuisce la trascrizione di *SNCA* in cellule HEK293T proteggendole dallo stress ossidativo (41). L'attività protettiva del miR-7 è stata anche dimostrata nei neuroni corticali primari trattati con MPP (N-metil-4-fenilpiridinio) dove la sovra-espressione di miR-7 ripristina la vitalità neuronale (42). I miR-7 e miR-153 sembrano avere un effetto sinergico determinando la riduzione dei livelli α Syn (43). Per quanto riguarda i *long non-coding* negli ultimi anni è stato identificato un NAT (Natural Antisense Transcript) del gene *SNCA*, chiamato *SNCA-AS1*. Ad oggi non sono ancora stati pubblicati studi sull'effetto di questo NAT e sul suo ruolo in condizioni fisiologiche o patologiche. A causa delle scarse informazioni sul ruolo di questo antisense e sull'azione di questa classe di lncRNA, è difficile fare un'ipotesi sulla possibile interazione tra l'antisense e il suo gene senso, anche se l'ipotesi più plausibile è che il primo vada a modulare l'espressione del secondo. Il NAT della proteina PINK (44) si compone di una porzione sovrapponibile al terminale 3' del gene codificante. Scheele et al. hanno dimostrato un'over-espressione di PINK1 a seguito di un'aumentata attività mitocondriale nel muscolo scheletrico e una correlazione nei livelli dei trascritti PINK1 e PINK1-AS (45, 46).

Recentemente è stato descritto un antisense naturale del gene *UCH-L1* (47,

48) che viene trascritto a partire dal secondo introne e che è sovrapponibile a questo solo per i primi 73 nucleotidi dell'mRNA del senso. In uno studio del 2012, è stato dimostrato che in condizioni fisiologiche il trascritto dell'antisense si localizza a livello nucleare, mentre in seguito a trattamento con Rapamicina trasloca nel citosol dei neuroni dopaminergici per andare ad interagire con l'mRNA di UCH-L1. L'accoppiamento dell'antisense con il suo gene senso porta a un aumento dei livelli di espressione di UCH-L1, suggerendo che tale aumento potrebbe rappresentare una risposta cellulare anti-apoptotica (49). Recentemente HOTAIR, un lncRNA antisense, è stato riportato come altamente espresso nel menencefalo di topi PD indotti da MPTP e in cellule di neuroblastoma umano (SH-SY5Y) trattate con MPP e sembrerebbe promuovere l'insorgenza di PD attraverso l'up-regolazione di LRRK2 (50). Dati preliminari hanno anche indicato che una up-regolazione di lncRNAp21, SNHG1 e MALAT1 precede l'insorgenza del PD (51). È stato inoltre dimostrato che MALAT1 si esprime in modo anomalo nei neuroni e modula una serie di geni associati allo sviluppo dendritico e sinaptico (52, 53). Inoltre, l'lncRNA nuclear paraspeckle assembly transcript 1 (NEAT1), una componente strutturale delle paraspeckles nucleari, regola positivamente il livello proteico di PINK1 inibendone la degradazione. NEAT1 promuove l'autofagia in un modello di PD indotto da MPTP attraverso un legame diretto con la proteina PINK1 (54).

Epigenetica nella malattia di Alzheimer

La malattia di Alzheimer (AD) è una malattia neurodegenerativa complessa per la quale sono stati identificati una serie di fattori di rischio come l'età, il sesso e i geni di suscettibilità alla patologia. Tuttavia, non è noto come tutti questi fattori influenzino l'insorgenza della patologia stessa. Ultimamente la ricerca si è quindi concentrata sullo studio delle variazioni epigenetiche al fine di identificare nuovi meccanismi eziologici. La *metilazione* del DNA (56-59) ha dimostrato alterazioni significative e riproducibili in quattro geni ANK1, RHBDF2, RPL13, e CDH23 (60-62). In particolare, il gene *ANK1*, che fino ad allora non era mai stato associato ad AD, mostra un'ampia regione di iper-metilazione su sei siti CpG a livello della corteccia cerebrale. Altre modificazioni epigenetiche che sembrano svolgere un ruolo chiave nell'AD sono le *modifiche istoniche* quali la tri-metilazione a livello della lisina 4 dell'istone H3, l'acetilazione della lisina 9,14,18 e 23 dell'istone H3 e della lisina 5, 8, 12 e 16 dell'istone H4, e la fosforilazione della serina 10 dell'istone H3. Al contrario la de-metilazione della lisina 9 dell'istone H3 è un tipico segno eterocromatico e quindi di repressione trascrizionale. Modifiche istoniche attivazionali sono state rilevate in varie regioni del cervello AD *post-mortem* come nella corteccia frontale (63, 64), nell'ippocampo (65-67), nella corteccia occipitale, mentre i cambiamenti trascrizionalmente repressivi sono stati osservati nel lobo temporale e occipitale (68), e nell'ippocampo (65). Studi su modelli animali hanno dimostrato che vi è una riduzione significativa dell'acetilazione degli istoni nei cervelli AD e ciò è associato a cambiamenti che possono essere revertiti utilizzando inibitori delle deacetilasi. Il fenibutirrato ad esempio revertirebbe qua-

si completamente la riduzione dell'acetilazione dell'istone H4 e permetterebbe il recupero dei difetti di memoria in topi mutati in APP (69). Risultati simili sono stati ottenuti anche con l'acido valproico (70, 71). Gli inibitori delle deacetilasi studiati nei modelli animali non sembrano avere effetto sulle cellule cerebrali umane (72) e ciò indurrebbe a pensare che gli inibitori delle deacetilasi abbiano un'azione specie-specifica.

Negli ultimi anni è sempre più evidente come anche l'alterazione nell'espressione dei *miRNA* contribuisca attivamente al processo di neurodegenerazione nell'AD (73, 74). I *miRNA* sembrano avere un ruolo nella regolazione dell'espressione di APP e di altri enzimi coinvolti nella produzione degli A β e principalmente della BACE1. Molti studi hanno dimostrato come vari *miRNA* regolino direttamente l'mRNA di APP in cellule umane *in vitro*. Questi includono miR-20a, miR-19 e miR-106a/b, miR-20 (75, 76) e miR-101 (77, 78) sia attraverso il legame diretto con il sito di legame posto a livello del promotore di APP (79), sia attraverso la regolazione dello splicing alternativo di APP (80-84). È stato dimostrato come i membri della famiglia dei miR-29 regolino l'espressione di BACE1 legandosi direttamente al suo 3'-UTR sia nelle linee di cellule umane sia di topo (85-87). Anche i *lncRNA* sono coinvolti nella patogenesi dell'AD. Uno dei *lncRNA* maggiormente studiato è l'antisense BACE1-AS che è trascritto dal filamento complementare della beta-secretasi-1 (88). L'espressione di BACE1-AS regola l'espressione della beta-secretasi stessa e questo determina l'incremento degli A β . Poche informazioni sono sinora disponibili sui *miRNA* e sul loro coinvolgimento nell'AD, inoltre l'utilizzo in terapia dei *miRNA* sono ancora in fase preliminare.

Epigenetica nella Sclerosi Laterale Amiotrofica

Nonostante sia stata clinicamente descritta 150 anni fa, i meccanismi responsabili della morte dei motoneuroni superiori e inferiori nella patogenesi della sclerosi laterale amiotrofica (SLA) (89-92) sono ad oggi sconosciuti. Alcuni studi indicano come le varie modifiche epigenetiche potrebbero far luce sui meccanismi che determinano la variabilità nell'insorgenza e la gravità della SLA sporadica aprendo così la strada all'identificazione di terapie efficaci e di strategie di diagnosi precoce. La metilazione del DNA è stata recentemente proposta come marker di disfunzione epigenetica nella SLA poiché essa aumenta nelle cellule del sangue indipendentemente dall'età di insorgenza della malattia (93). Inoltre, in campioni di midollo spinale di SLA sporadica è stato riportato un aumento sia dei livelli di 5mC sia dei livelli 5hmC (94). Un chiaro esempio di come i cambiamenti epigenetici possano modificare l'espressione genica è rappresentato dalle mutazioni a carico di CORF72 (95). Queste espansioni esanucleotidiche localizzate nella regione promotrice del gene sono associate a molteplici meccanismi patogenici, tra i quali l'acquisizione di citotossicità degli RNA mutati e la successiva formazione di ibridi di RNA-DNA stabili (R-loop) (96). Diversi gruppi hanno segnalato la presenza di ipermetilazione delle isole CpG situate nel promotore C9ORF72 in circa il 10-30% dei pazienti affetti da FTD-ALS C9ORF72 (97-100). Per quanto riguarda le *modifiche istoniche*, uno studio del 2003 riportava una forte dimi-

nuzione della forma acetilata dell'istone H3 in un modello di topo SLA SOD1 mutato in particolare nel nucleo del motoneurone (101). In campioni post-mortem di corteccia motoria e midollo spinale di pazienti SLA è stata inoltre trovata una sovra-regolazione dei livelli di mRNA HDAC2 e una concomitante riduzione nei livelli di mRNA HDAC11 (102). In una coorte di pazienti SLA, a decorso rapido di malattia, mostravano un aumento del livello di espressione di HDAC4 nel muscolo scheletrico rispetto ai pazienti con aspettativa di vita più lunga (103-105). La biogenesi dei *miRNA* potrebbe essere modificata dall'interazione di TDP-43 con Drosha (enzima deputato al clivaggio dei precursori dei miRNA) creando un complesso alternativo che faciliterebbe la sintesi di diversi pre-miRNA (106). Il silenziamento di TDP-43, in linee cellulari di linfoblasti umani di pazienti SLA, determina una ridotta espressione di diversi miRNA (miR-132-5p/3p, miR-143-5/3p e miR-574-3p) (107). Tuttavia, nella SLA è stata segnalata anche la sovra-regolazione di diversi miRNA, come ad esempio l'over-espressione del miR-155 in modelli murini SOD1-G93A, ulteriormente confermata nel midollo spinale SLA umano (108, 109). Esperimenti con inibitori di miR-155 hanno mostrato un aumento significativo del tasso di sopravvivenza nei topi SLA (109). Anche la disfunzione mitocondriale del muscolo scheletrico è tipica della SLA e, anche in questo tessuto, è stata riscontrata una maggiore espressione di specifici miRNA, tra cui miR-23a, miR-29b, miR-206 e miR-455. Ci sono diverse funzioni attribuite a questi miRNA; miR-23a sopprime i livelli di mRNA del PGC-1 α , un fattore di trascrizione coinvolto nella termogenesi e nell'adattamento muscolare all'esercizio fisico, miR-29b è coinvolto nella rigenerazione muscolare e il miR-455 ha un ruolo nella degenerazione muscolare (110). Ad oggi vi sono pochi dati relativi ai *lncRNA* nella SLA, tuttavia evidenze sempre maggiori sottolineano come l'alterazione del metabolismo dell'RNA giochi un ruolo importante nella patogenesi della malattia. Ciò è supportato anche dal fatto che mutazioni tipiche della SLA siano a carico di geni come TARDBP e FUS che codificano per proteine che legano DNA/RNA (DNA/RNA binding proteins) e che sono coinvolte nella trascrizione e nel processamento dell'RNA, inclusa la biogenesi del miRNA (111). Analogamente *SOD1*, *Alsin* e *ANG* sono coinvolti nella trascrizione e nell'elaborazione degli RNA codificanti e non codificanti (112) e SOD1 mutata induce una down-regolazione dello splicing alternativo (111). Anche se SOD1, TDP-43 e FUS sono comunemente coinvolti nella SLA attraverso percorsi differenti (113), la loro comune implicazione nell'omeostasi dell'espressione genica suggerisce che alterazioni della regolazione dell'espressione possono rappresentare eventi chiave nella patogenesi della SLA. Finora, è stato riportato che NEAT1 genera due tipi di RNA lunghi non codificanti, uno di questi, denominato NEAT1_2 lncRNA che interagisce con la formazione di paraspeckle nei neuroni motori spinali di pazienti affetti da SLA¹¹⁴. Infatti, è stato dimostrato che NEAT1_2 lncRNA è sovra-regolato durante la fase iniziale della patogenesi della SLA rispetto ai controlli sani (114). Inoltre, nei pazienti affetti da demenza frontotemporale, sebbene l'interazione meccanica tra lncRNA e RNA codificanti non sia ancora del tutto chiara, sembrerebbe che TDP-43 e FUS si leghino a diversi lncRNA regolandone l'espressione (111, 115).

Bibliografia

1. Waddington CH. The epigenotype. 1942. *Int. J. Epidemiol.* 2012; 41: 10-13.
2. Sweatt JD. The emerging field of neuroepigenetics. *Neuron.* 2013; 80: 624-632.
3. Maze I, Noh KM, Soshnev AA, et al. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nat. Rev. Genet.* 2014; 15: 259-271.
4. Piletic K, Kunej T. MicroRNA epigenetic signatures in human disease. *Arch. Toxicol.* 2016; 90: 2405-2419.
5. Saeidimehr S, Ebrahimi A, Saki N, et al. MicroRNA-based linkage between aging and cancer: from epigenetics view point. *Cell J.* 2016; 18: 117-126.
6. Leone S, Santoro R. Challenges in the analysis of long noncoding RNA functionality. *FEBS Lett.* 2016; 590: 2342-2353.
7. Roberts TC, Morris KV, Wood MJ. The role of long non-coding RNAs in neurodevelopment, brain function and neurological disease. *Phil. Trans. R. Soc. B.* 2014; 369: 20130507.
8. Feng Y, Jankovic J, Wu YC. Epigenetic mechanisms in Parkinson's disease. *J Neurol Sci.* 2015; 349: 3-9.
9. Paquet D, Kwart D, Chen A, et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature.* 2016; 533: 125-129.
10. Xue HY, Ji LJ, Gao AM, et al. CRISPR-Cas9 for medical genetic screens: applications and future perspectives. *J Med Genet.* 2016; 53: 91-97.
11. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science.* 2009; 324: 930-935.
12. Ficz G, Branco MR, Seisenberger S, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature.* 2011; 473: 398-402.
13. Branco MR, Ficz G, Reik W. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet.* 2011; 13: 7-13.
14. Urduinguo RG, Sanchez-Mut JV, Esteller M. Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. *Lancet Neurol.* 2009; 8: 1056-1072.
15. Iyer MK, Niknafs YS, Malik R, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet.* 2015; 47: 199-208.
16. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell.* 2009; 136: 215-233.
17. Cabili MN, Trapnell C, Goff L, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 2011; 25: 1915-1927.
18. Salta E, De Strooper B. Non-coding RNAs with essential roles in neurodegenerative disorders. *Lancet Neurol.* 2012; 11: 189-200.
19. Qureshi IA, Mehler MF. Long non-coding RNAs: novel targets for nervous

- system disease diagnosis and therapy. *Neurotherapeutics*. 2013; 10: 632-646.
20. Kraus TF, Haider M, Spanner J, et al. Altered long noncoding RNA expression precedes the course of Parkinson's disease - a preliminary report. *Mol Neurobiol*. 2016; 8: 1-9.
 21. Kapusta Z, Kronenberg VJ, Lynch X, et al. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet*. 2011; 9: e1003470.
 22. Bernstein E, Allis CD. RNA meets chromatin. *Genes Dev*. 2005; 19: 1635-1655.
 23. Whitehead J, Pandey GK & Kanduri C. Regulation of the mammalian epigenome by long noncoding RNAs. *Biochim. Biophys. Acta*. 2009; 1790: 936-947.
 24. Pollard KS, Salama SR, Lambert N, et al. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature*. 2006; 443: 167-172.
 25. Hashimoto-Torii K, Motoyama J, Hui CC, et al. Differential activities of Sonic hedgehog mediated by Gli transcription factors define distinct neuronal subtypes in the dorsal thalamus. *Mech. Dev*. 2003; 120: 1097-1111.
 26. Masliah E, Dumaop W, Galasko D, et al. Distinctive patterns of DNA methylation associated with Parkinson disease: identification of concordant epigenetic changes in brain and peripheral blood leukocytes. *Epigenetics*. 2013; 8: 1030-1038.
 27. Moore K, McKnight AJ, Craig D, et al. Epigenome-wide association study for Parkinson's disease. *Neuromolecular Med*. 2014; 16: 845-855.
 28. Desplats P, Spencer B, Coffee E, et al. Alpha-synuclein sequesters Dnmt1 from the nucleus: a novel mechanism for epigenetic alterations in Lewy body diseases. *J Biol Chem*. 2011; 286: 9031-9037.
 29. Kaut O, Schmitt I, Wüllner U. Genome-scale methylation analysis of Parkinson's disease patients' brains reveals DNA hypomethylation and increased mRNA expression of cytochrome P450 2E1. *Neurogenetics*. 2012; 13: 87-91.
 30. Cai M, Tian J, Zhao G, et al. Study of methylation levels of parkin gene promoter in Parkinson's disease patients. *Int J Neurosci*. 2011; 121: 497-502.
 31. Kagara I, Enokida H, Kawakami K, et al. CpG Hypermethylation of the UCHL1 gene promoter is associated with pathogenesis and poor prognosis in renal cell carcinoma. *J Urol*. 2008; 180: 343-351.
 32. Behrens MI, Brüggemann N, Chana P, et al. Clinical spectrum of Kufor-Rakeb syndrome in the Chilean kindred with ATP13A2 mutations. *Mov Disord*. 2010; 25: 1929-1937.
 33. van Heesbeen HJ, Mesman S, Veenliet JV, et al. Epigenetic mechanisms in the development and maintenance of dopaminergic neurons. *Development*. 2013; 140: 1159-1169.
 34. Park G, Tan J, Garcia G, et al. Regulation of Histone Acetylation by Autophagy in Parkinson Disease. *J Biol Chem*. 2016; 291: 3531-3540.
 35. Song C, Kanthasamy A, Jin H, et al. Paraquat induces epigenetic changes by promoting histone acetylation in cell culture models of dopaminergic degeneration. *Neurotoxicology*. 2011; 32: 586-595.

36. Goers J, Manning-Bog AB, McCormack AL, et al. Nuclear localization of alpha-synuclein and its interaction with histones. *Biochemistry*. 2003; 42: 8465-8471.
37. Jankovic J, Chen S, Le WD. The role of Nurr1 in the development of dopaminergic neurons and Parkinson's disease. *Prog Neurobiol*. 2005; 77: 128-138.
38. Lungu G, Stoica G, Ambrus A. MicroRNA profiling and the role of microRNA-132 in neurodegeneration using a rat model. *Neurosci Lett*. 2013; 553: 153-158.
39. Khoo SK, Petillo D, Kang UJ, et al. Plasma-based circulating microRNA biomarkers for Parkinson's disease. *J Park Dis*. 2012; 2: 321-331.
40. Cho HJ, Liu G, Jin SM, et al. MicroRNA-205 regulates the expression of Parkinson's disease-related leucine-rich repeat kinase 2 protein. *Hum Mol Genet*. 2013; 22(3): 608-620.
41. Junn E, Lee KW, Jeong BS, et al. Repression of alpha-synuclein expression and toxicity by microRNA-7. *Proc Natl Acad Sci USA*. 2009; 106: 13052-13057.
42. Fragkouli A, Doxakis E. miR-7 and miR-153 protect neurons against MP-P(+)-induced cell death via upregulation of mTOR pathway. *Front Cell Neurosci*. 2014; 8: 182.
43. Doxakis E. Post-transcriptional regulation of alpha-synuclein expression by mir-7 and mir-153. *J Biol Chem*. 2010; 285: 12726-12734.
44. Blackinton JG, Anvret A, Beilina A, et al. Expression of PINK1 mRNA in human and rodent brain and in Parkinson's disease. *Brain Res*. 2007; 118: 10-16
45. Scheele C, Petrovic N, Faghihi MA, et al. The human PINK1 locus is regulated in vivo by a non-coding natural antisense RNA during modulation of mitochondrial function. *BMC Genomics*. 2007; 8: 74.
46. Chiba M, Kiyosawa H, Hiraiwa N, et al. Existence of Pink1 antisense RNAs in mouse and their localization. *Cytogen Genome Res*. 2009; 126: 259-270.
47. Maraganore DM, Lesnick TG, Elbaz A, et al. UCHL1 is a Parkinson's disease susceptibility gene. *Ann Neurol*. 2004; 55: 512-521.
48. Barrachina M, Castano E, Dalfo E, et al. Reduced ubiquitin C-terminal hydrolase-1 expression levels in dementia with Lewy bodies. *Neurobiol. Dis*. 2006; 22: 265-273.
49. Carrieri C, Cimatti L, Biagioli M, et al. Long non-coding antisense RNA controls Uchl1 translation through an embedded INEB2 repeat. *Nature*. 2012; 491: 454-457.
50. Liu S, Cui B, Dai ZX, et al. Long non-coding RNA HOTAIR promotes Parkinson's disease induced by MPTP through up-regulating the expression of LRRK2. *Curr Neurovasc Res*. 2016; 13: 115-120.
51. Kraus TF, Haider M, Spanner J, et al. Altered long noncoding RNA expression precedes the course of Parkinson's disease - a preliminary report. *Mol Neurobiol*. 2016; 8: 1.
52. Bernard D, Prasanth KV, Tripathi V, et al. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *EMBO J*. 2010; 29: 3082-3093.

53. Lipovich L, Dachet F, Cai J, et al. Activity-dependent human brain coding/noncoding gene regulatory networks. *Genetics* 2012; 192: 1133-1148.
54. Yan W, Chen ZY, Chen JQ, et al. LncRNA NEAT1 promotes autophagy in MPTP-induced Parkinson's disease through stabilizing PINK1 protein. *Biochem Biophys Res Commun.* 2018; 496: 1019-1024.
55. Brookmeyer R, Johnson E, Ziegler-Graham K, et al. Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement.* 2007; 3: 186-191.
56. Song CX, Yi C, He C. Mapping recently identified nucleotide variants in the genome and transcriptome. *Nat Biotechnol.* 2012; 30: 1107-1116.
57. Stewart SK, Morris TJ, Guilhamon P, et al. oxBS-450 K: a method for analysing hydroxymethylation using 450 K BeadChips. *Methods.* 2015; 72: 9-15.
58. Lunnon K, Hannon E, Smith RG, et al. Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. *Genome Biol.* 2016; 17: 27.
59. Field SF, Beraldi D, Bachman M, et al. Accurate measurement of 5-methylcytosine and 5-hydroxymethylcytosine in human cerebellum DNA by oxidative bisulfite on an array (OxBS-array). *PLoS One.* 2015; 10: e0118202.
60. De Jager PL, Srivastava G, Lunnon K, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat Neurosci.* 2014; 17: 1156-1163.
61. Lunnon K, Smith R, Hannon E, et al. Methylomic profiling implicates cortical de-regulation of ANK1 in Alzheimer's disease. *Nat Neurosci.* 2014; 17: 1164-1170.
62. Lord J, Cruchaga C. The epigenetic landscape of Alzheimer's disease. *Nat Neurosci.* 2014; 17: 1138-1140.
63. Rao JS, Keleshian VL, Klein S, et al. Epigenetic modifications in frontal cortex from Alzheimer's disease and bipolar disorder patients. *Transl Psychiatry.* 2012; 2: e132.
64. Lu T, Aron L, Zullo J, et al. REST and stress resistance in ageing and Alzheimer's disease. *Nature.* 2014; 507: 448-454.
65. Hernández-Ortega K, Garcia-Esparcia P, Gil L, et al. Altered machinery of protein synthesis in Alzheimer's: from the nucleolus to the ribosome. *Brain Pathol.* 2015; 26: 593-605.
66. Mastroeni D, Delvaux E, Nolz J, et al. Aberrant intracellular localization of H3k4me3 demonstrates an early epigenetic phenomenon in Alzheimer's disease. *Neurobiol Aging.* 2015; 36: 3121-3129.
67. Ogawa O, Zhu X, Lee HG, et al. Ectopic localization of phosphorylated histone H3 in Alzheimer's disease: a mitotic catastrophe? *Acta Neuropathol.* 2003; 105: 524-528.
68. Zhang K, Schrag M, Crofton A, et al. Targeted proteomics for quantification of histone acetylation in Alzheimer's disease. *Proteomics.* 2012; 12: 1261-1268.
69. Ricobaraza A, Cuadrado-Tejedor M, Perez-Mediavilla A, et al. Phenylbutyrate ameliorates cognitive deficit and reduces tau pathology in an Alzheimer's disease mouse model. *Neuropsychopharmacology.* 2009; 34: 1721-172.
70. Yao ZG, Liang L, Liu Y, et al. Valproate improves memory deficits in an Alzheimer's disease mouse model: investigation of possible mechanisms of action. *Cell Mol Neurobiol.* 2014; 34: 805-812.

71. Qing H, He G, Ly PTT, et al. Valproic acid inhibits $a\beta$ production, neuritic plaque formation, and behavioral deficits in Alzheimer's disease mouse models. *J Exp Med*. 2008; 205: 2781-2789.
72. Gibbons HM, Smith AM, Teoh HH, et al. Valproic acid induces microglial dysfunction, not apoptosis, in human glial cultures. *Neurobiol Dis*. 2011; 41: 96-103.
73. Hardy JA & Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. *Science*. 1992; 256: 184.
74. Karran E, Mercken M, De Strooper B. The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nat Rev Drug Discov*. 2011; 10: 698-712
75. Fan X, Liu Y, Jiang J, et al. miR-20a promotes proliferation and invasion by targeting APP in human ovarian cancer cells. *Acta Biochim Biophys Sin*. 2010; 42: 318-324.
76. Hebert SS, Horre K, Nicolai L, et al. MicroRNA regulation of Alzheimer's amyloid precursor protein expression. *Neurobiol Dis*. 2009; 33: 422-428.
77. Vilardo E, Barbato C, Ciotti M, et al. MicroRNA-101 regulates amyloid precursor protein expression in hippocampal neurons. *J Biol Chem*. 2010; 285: 18344-18351.
78. Long JM, Ray B, Lahiri DK. MicroRNA-339-5p down-regulates protein expression of beta site amyloid precursor protein-cleaving enzyme 1 (BACE1) in human primary brain cultures and is reduced in brain tissue specimens of Alzheimer disease subjects. *J Biol Chem*. 2014; 289: 5184-5198.
79. Delay C, Calon F, Mathews P, et al. Alzheimer-specific variants in the 3'UTR of Amyloid precursor protein affect microRNA function. *Mol Neurodegener*. 2011; 6: 70.
80. Donev R, Newall A, Thome J, et al. A role for SC35 and hnRNPA1 in the determination of amyloid precursor protein isoforms. *Mol Psychiatry*. 2007; 12: 681-690.
81. Smith P, Al Hashimi A, Girard J, et al. In vivo regulation of amyloid precursor protein neuronal splicing by microRNAs. *J Neurochem*. 2011; 116: 240-247.
82. Lukiw WJ. Micro-RNA speciation in fetal, adult and Alzheimer's disease hippocampus. *Neuroreport*. 2007; 18: 297-300.
83. Kong Y, Wu J, Zhang D, et al. The role of miR-124 in Drosophila Alzheimer's disease model by targeting delta in notch signaling pathway. *Curr Mol Med*. 2015; 15: 980-989.
84. Schonrock N, Matamales M, Ittner LM, et al. MicroRNA networks surrounding APP and amyloid- β metabolism - implications for Alzheimer's disease. *Exp Neurol*. 2012; 235: 447-454.
85. Yang G, Song Y, Zhou X, et al. MicroRNA-29c targets beta-site amyloid precursor protein-cleaving enzyme 1 and has a neuroprotective role in vitro and in vivo. *Mol Med Rep*. 2015; 12: 3081-3088.
86. Lei X, Lei L, Zhang Z, et al. Downregulated miR-29c correlates with increased BACE1 expression in sporadic Alzheimer's disease. *Int J Clin Exp Pathol*. 2015; 8: 1565-1574.

87. Zong Y, Wang H, Dong W, et al. miR-29c regulates BACE1 protein expression. *Brain Res.* 2011; 1395: 108-115.
88. Faghihi MA, Modarresi F, Khalil AM, et al. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of [beta]-secretase. *Nat Med.* 2008; 14: 723-730.
89. Dangoumau A, Verschuere A, Hammouche E, et al. Novel SOD1 mutation p.V31A identified with a slowly progressive form of amyotrophic lateral sclerosis. *Neurobiol Aging.* 2014; 35: 266 e1-4.
90. Forsberg K, Andersen PM, Marklund SL, et al. Glial nuclear aggregates of superoxide dismutase-1 are regularly present in patients with amyotrophic lateral sclerosis. *Acta Neuropathol.* 2011; 121: 623-634.
91. Ivanova MI, Sievers SA, Guenther EL, et al. Aggregation-triggering segments of SOD1 fibril formation support a common pathway for familial and sporadic ALS. *Proc Natl Acad Sci USA.* 2014; 111: 197-201.
92. Majounie E, Renton AE, Mok K, et al. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* 2012; 11: 323-330.
93. Tremolizzo L, Messina P, Conti E, et al. Whole-blood global DNA methylation is increased in amyotrophic lateral sclerosis independently of age of onset. *Amyotroph Lateral Scler Frontotemporal Degener.* 2014; 15: 98-105.
94. Figueroa-Romero C, Hur J, Bender DE, et al. Identification of epigenetically altered genes in sporadic amyotrophic lateral sclerosis. *PLoS One.* 2012; 7: e52672.
95. Caballero-Hernandez D, Toscano MG, Cejudo-Guillen M, et al. The 'Omics' of amyotrophic lateral sclerosis. *Trends Mol Med.* 2016; 22: 53-67.
96. He F & Todd PK. Epigenetics in nucleotide repeat expansion disorders. *Semin Neurol.* 2011; 31: 470-483.
97. Belzil VV, Bauer PO, Gendron TF, et al. Characterization of DNA hypermethylation in the cerebellum of c9FTD/ALS patients. *Brain Res.* 2014; 1584: 15-21.
98. Liu EY, Russ J, Wu K, et al. C9orf72 hypermethylation protects against repeat expansion-associated pathology in ALS/FTD. *Acta Neuropathol.* 2014; 128: 525-541.
99. Xi Z, Rainero I, Rubino E, et al. Hypermethylation of the CpG-island near the C9orf72 G(4)C(2)-repeat expansion in FTLD patients. *Hum Mol Genet.* 2014; 23: 5630-5637.
100. Xi Z, Zinman L, Moreno D, et al. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *Am J Hum Genet.* 2013; 92: 981-989.
101. Rouaux C, Jokic N, Mbebi C, et al. Critical loss of CBP/p300 histone acetylase activity by caspase-6 during neurodegeneration. *EMBO J.* 2003; 22: 6537-6549.
102. Janssen C, Schmalbach S, Boeselt S, et al. Differential histone deacetylase mRNA expression patterns in amyotrophic lateral sclerosis. *J Neuropathol Exp Neurol.* 2010; 69: 573-581.

103. Bruneteau G, Simonet T, Bauche S, et al. Muscle histone deacetylase 4 up-regulation in amyotrophic lateral sclerosis: potential role in reinnervation ability and disease progression. *Brain*. 2013; 136: 2359-2368.
104. Rouaux C, Panteleeva I, Rene F, et al. Sodium valproate exerts neuroprotective effects in vivo through CREB-binding protein-dependent mechanisms but does not improve survival in an amyotrophic lateral sclerosis mouse model. *J Neurosci*. 2007; 27: 5535-5545.
105. Petri S, Kiaei M, Kipiani K, et al. Additive neuroprotective effects of a histone deacetylase inhibitor and a catalytic antioxidant in a transgenic mouse model of amyotrophic lateral sclerosis. *Neurobiol Dis*. 2006; 22: 40-49.
106. Kawahara Y & Mieda-Sato A. TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc Natl Acad Sci USA*. 2012; 109: 3347-3352.
107. Freischmidt A, Muller K, Ludolph AC, et al. Systemic dysregulation of TDP-43 binding microRNAs in amyotrophic lateral sclerosis. *Acta Neuropathol Commun*. 2013; 1: 42.
108. Butovsky O, Siddiqui S, Gabriely G, et al. Modulating inflammatory monocytes with a unique microRNA gene signature ameliorates murine ALS. *J Clin Invest*. 2012; 122: 3063-3087.
109. Koval ED, Shaner C, Zhang P, et al. Method for widespread microRNA-155 inhibition prolongs survival in ALS-model mice. *Hum Mol Genet*. 2013; 22: 4127-4135.
110. Russell AP, Wada S, Vergani L, et al. Disruption of skeletal muscle mitochondrial network genes and miRNAs in amyotrophic lateral sclerosis. *Neurobiol Dis*. 2013; 49: 107-117.
111. Strong, MJ. The evidence for altered RNA metabolism in amyotrophic lateral sclerosis (ALS). *J Neurol Sci* 2010; 288:1-12.
112. Milani P, Amadio M, Laforenza U, et al. Posttranscriptional regulation of SOD1 gene expression under oxidative stress: Potential role of ELAV proteins in sporadic ALS. *Neurobiol Dis*. 2013; 60: 51-60.
113. Chen Y, Wei Q, Chen X, et al. Aberration of miRNAs expression in leukocytes from sporadic amyotrophic lateral sclerosis. *Front Mol Neurosci*. 2016; 9: 69.
114. Nishimoto Y. The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. *Mol Brain*. 2013; 6: 31.
115. Lourenco GF, Janitz M, Huang Y, et al. Long noncoding RNAs in TDP-43 and FUS/TLS-related frontotemporal lobar degeneration (FTLD). *Neurobiol Dis*. 2015; 82: 445-454.

Epigenetics, obesity and diabetes

Lucia Migliore

Department of Traslational Research and New Technologies in Medicine and Surgery, University of Pisa

Currently there is a growing interest in obesity, not only for its increasing prevalence worldwide, but also because it can be the cause of many other pathologies, including type 2 diabetes (T2D), cardiovascular disease, asthma, liver and pulmonary problems, gall bladder and kidney disease, osteoarthritis, reproductive problems and many types of cancer. Indeed type 2 diabetes, previously considered a disease with adult onset, has risen sharply in children and young people simultaneously with the increase in obesity. To state that the main cause of obesity is to be found in an imbalance between energy intake and expenditure due the current adoption of a sedentary lifestyle, coupled with the easier access to high-caloric food, is, in the light of the most recent discoveries, too simplistic. The genetic contribution to obesity and diabetes has been demonstrated by means of linkage analysis, twin, and adoption studies. Eight well-established monogenic obesity genes have been identified: leptin (*LEP*), leptin receptor (*LEPR*), proopiomelanocortin (*POMC*), prohormone convertase 1 (*PCSK1*), melanocortin 4 receptor (*MC4R*), single-minded homologue 1 (*SIMI*), brain-derived neurotrophic factor (*BDNF*), and neurotrophic tyrosine kinase receptor type 2 (*NTRK2*). Mutations in these genes are causative of early onset forms of obesity and hyperphagia and may account for about 10% of severely obese children (1). Apart from the rare monogenic forms of obesity, most cases of obesity are sporadic, so they are complex forms, where both predisposition genes and environmental factors are likely involved. Genome-wide association studies (GWAS) have identified a high number of genetic polymorphisms associated with complex human diseases and traits, and have provided valuable information into their genetic structure. Many regions in the genome have been associated with obesity and T2D; in general we can say that hundreds of loci are involved in both common and rare variants. The first single nucleotide polymorphism (SNP) identified in 2007 as significantly associated with increased body mass index (BMI) was found in the *FTO* gene and resulted to be associated to fat mass and obesity and to affect obesity by regulating appetite (2). By searching the medical literature and computer generated websites for key words, Butler and co-workers in 2015 documented evidence for 370 genes playing a role in obesity. The obesity genes primarily affected common pathways in lipid metabolism, deposition or transport, eating behavior and food selection, physical activity or energy expenditure. Twenty-one of the obesity genes were also associated with human infertility and reproduction (3). The last review on this topic sum-

marizes a total of 72 studies, resulting in 90,361 susceptibility single nucleotide polymorphisms for obesity, of which over 90% are located in non-coding regions (4). However the majority of variants identified confer relatively small increments in risk, or even in protection, and explain only a small portion of familial clustering, leaving the question of “missing” heritability” still unexplained. To explain the origin of pandemic of obesity of the last decades from an evolutionary point of view, some hypotheses were advanced, such as the “*thrifty genotype hypothesis*” by J. Neel (5) or “*thrifty phenotype hypothesis*” by J.V. Hales (6). However, these hypotheses have been well synthesised in the following years, when in the light of epigenetic studies, it became evident that the entire embryo-foetal and perinatal period of development plays a key role in the programming of all human organs and tissues. A new and general pathogenic paradigm, the Developmental Origins of Health and Disease (DOHaD) theory was proposed (7), to better explain the current epidemiological transition, consisting of the worldwide increase of chronic, degenerative and inflammatory diseases such as obesity, diabetes, cardiovascular, neurodegenerative disease and cancer. It is now emerging that obesity and its related complications are associated with other factors such as environmental pollutants (obesogens), gut microbiota and specific nutrients, which can increase susceptibility to weight gain and to other metabolic consequences through epigenetic changes (8). The most important epigenetic mechanisms include DNA methylation, histone tail modifications and non coding RNAs (ncRNAs) interventions.

An example of the importance of maternal exposure (diet) is provided by the classical study on the Avy allele at the Agouti locus in mice. Maternal supplementation during pregnancy with donors of methyl groups can modify the phenotype of the progeny by interfering with the expression of the agouti locus responsible for the colour of the coat (and other phenotypic traits, including body weight) (9). Studies focused on epigenetic marks in obesity found altered methylation and/or histone acetylation levels in genes involved in specific but also in more general metabolic processes (10).

Analogously changes in epigenetic marks and differential regulation of epigenetic modulators have been observed in different models of diabetes and its associated complications (including case-control cohorts, prospective cohorts, intervention studies, and cultures of human cells exposed to environmental risk factors) (11, 12).

In 2006 Grün and Blumberg formulated the obesogen hypothesis, according to which obesity epidemic would be at least in part consequence of the spread in the environment, especially within food chain, of xenobiotics able to act as endocrine disruptors (EDC), mainly during foetal programming, by promoting hyperplasia of the adipocyte pool, facilitating adipogenic pathways that activate the hyperplasia during periods of increased physiological development, perturbing the lipid homeostasis, interfering with feedback mechanisms of appetite and satiety (13). Among obesogens, which can be defined from a functional point of view as chemicals that, improperly interfering with lipid homeostasis, promote adipogenesis, there are bisphenol A, organochlorine pesticides (DDT, DDE...), polychlorobiphenils, brominated flame retardant, dioxins, phtalates, diethylstilbestrol (DES), heavy metals

(arsenic, cadmium, lead...). Little is known about epigenetic changes induced by obesogenic exposures. As for the possible epigenetic and transgenerational effects of EDC, the most studied compound is bisphenol A (BPA), a synthetic chemical found in food and beverage containers, baby bottles, and dental materials. On the Agouti mouse model BPA was found to induce hypomethylation and increased expression of the Agouti gene in prenatally exposed mice (early developmental stages generally represent the period of greatest sensitivity to these chemicals, see figure 1) led to the birth of mice characterized by yellow rather than brown fur, as well as by tendency to develop obesity, diabetes and tumors (14). It is now clear from studies in animal and human models that there are effects of maternal obesity on the progeny, and that these act even before birth. Altered methylation outcomes at multiple imprint regulatory regions were found in children born to obese parents, compared with children born to non-obese parents. Moreover it seems that there is a preconceptional influence by parental life-style or overnutrition on the (re)programming of imprint marks during gametogenesis (both male and female) and early development (15). Although obesity can be passed on to offspring via genetic variants, there is growing evidence that prenatal epigenetic programming plays a substantial role in the transmission of obesity across generations, according to the hypothesis that epigenetic modifications can be passed to the successive generations (intergenerational and transgenerational inheritance) (Figure 1).

Thanks to studies in the field of epigenetics, molecular mechanisms underlying many pathogenic pathways are starting to be elucidated. There is evidence that many nutritional factors could act by epigenetic mechanisms, modulating DNA methylation or histone modifications and some of them might be used successfully in obesity therapy. Some good results have already been achieved through the use of the methyl donors (folate, methionine, choline and vitamin B12), especially by

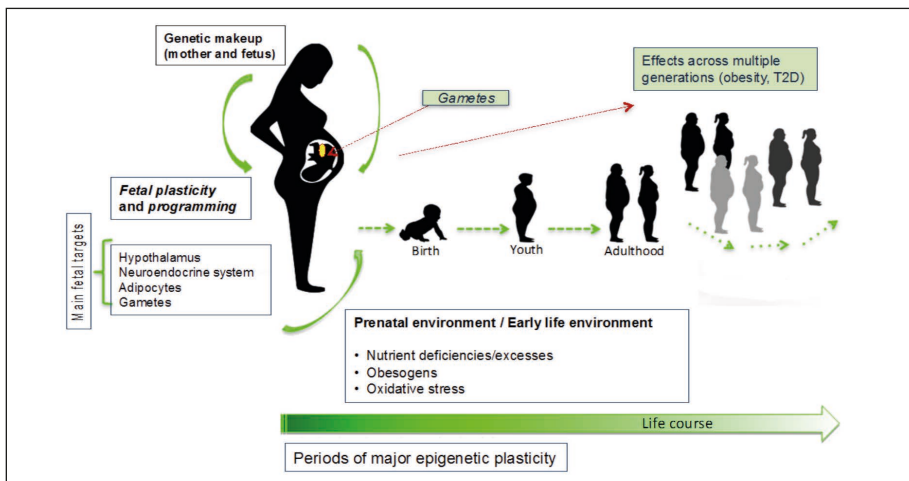


Fig. 1 - Importance of environmentally driven epigenetic effects for obesity and type 2 diabetes during life course and potential consequences across generations, according to the DOHaD theory. (From Burgio et al., 2015).

maternal diet supplementation (8). Recently, many epigenetic biomarkers have been identified linked to a greater risk of developing complex diseases, including obesity and T2D. Such epigenetic signatures, if confirmed in studies on large populations and if found to have high sensitivity and specificity, could help to improve the diagnosis, the prognosis and to introduce new potential therapeutical interventions.

Bibliografia

1. Loos RJ. Genetic determinants of common obesity and their value in prediction. *Best Pract Res Clin Endocrinol Metab.* 2012; 26: 211-226.
2. Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet.* 2007; 8: 657-662.
3. Butler MG, McGuire A, Manzardo AM. Clinically relevant known and candidate genes for obesity and their overlap with human infertility and reproduction. *J Assist Reprod Genet.* 2015; 32: 495-508.
4. Dong S-S, Zhang Y-J, Chen Y-X, et al. Comprehensive review and annotation of susceptibility SNPs associated with obesity-related traits. *Obes Rev.* 2018 Mar 12. doi: 10.1111/obr.12677
5. Neel JV. Diabetes mellitus: A “thrifty” genotype rendered detrimental by “progress”? *Am J Hum Genet.* 1962; 14: 353-352
6. Hales CN, Barker DJ. The thrifty phenotype hypothesis. *Br. Med. Bull.* (2001) 60: 5-20.
7. Barker DJ, Eriksson JG, Forsén T, Osmond C. Fetal origins of adult disease: strength of effects and biological basis. *Int J Epidemiol.* 2002; 31: 1235-1239.
8. Burgio E, Lopomo A, Migliore L. Obesity and diabetes: from genetics to epigenetics. *Mol Biol Rep.* 2015; 42: 799-818.
9. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol.* 2003; 23: 5293-5300.
10. Sayols-Baixeras S, Subirana I, Fernandez-Sanles A, et al. DNA methylation and obesity traits: An epigenome-wide association study. *The REGICOR study. Epigenetics.* 2017, 12: 909-916.
11. Al-Haddad, R, Karniba N, Assaada RA, et al. Epigenetic changes in diabetes. *Neuroscience Letters* 2016; 625 64-69.
12. Davegårdh C, García-Calzón S, Bacos K, Ling C. DNA methylation in the pathogenesis of type 2 diabetes in humans. *Molecular Metabolism.* 2018. doi: 10.1016/j.molmet.2018.01.022
13. Grün F, Blumberg B. Environmental obesogens: organotins and endocrine disruption via nuclear receptor signaling. *Endocrinology.* 2006; 147 (Suppl. 6): S50-55.
14. Dolinoy DC. The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. *Nutr Rev.* 2008; 66 (Suppl. 1): S7-11.
15. Soubry A, Murphy SK, Wang F, et al. Newborns of obese parents have altered DNA methylation patterns at imprinted genes. *Int J Obes (Lond).* 2015; 39: 650-657.

Epigenetic mechanism in early mammalian development

Maria Elena Torres-Padilla

Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum, München, Germany

In mammals, the terminally differentiated sperm and oocyte fuse to create a totipotent zygote upon fertilisation, which is not lineage-restricted and has the ability to generate all cell types in a new organism. The zygote undergoes five cell divisions to give rise to the blastocyst, which is composed of two distinct lineages: the inner cell mass (ICM) that gives rise to the embryo proper, and the trophectoderm that mainly gives rise to the placental tissue. Embryonic stem cells (ESCs) derive from the ICM and thus they are pluripotent but not totipotent. However, because ESCs can be stably maintained *in vitro* and because of their potential for regenerative medicine, the transcriptional network and epigenetic status governing pluripotency in ESCs have been extensively studied. These efforts led to the discovery of the generation of induced pluripotent stem (iPS) cells; which requires specific transcription factors previously identified as master regulators of pluripotency such as Oct3/4 (Pou5f1) (1). Addition of this specific ‘cocktail’ of transcription factors to somatic cells from the three major lineages can generate pluripotent cells. Perhaps owing to the lack of knowledge about the molecular regulatory mechanisms governing totipotency, it remains unclear if a totipotent cell that is a cell with an even greater degree of plasticity, can be maintained *in vitro* or induced experimentally as in the case of iPS cells. Currently, the only way to generate a totipotent cell experimentally is through somatic cell nuclear transfer (SCNT), which involves the physical transplantation of a somatic nucleus into a nucleated oocyte. SCNT has recently been used in humans to derive embryonic stem cells (2). However, because SCNT has a low success rate and requires using oocytes, there have been both scientific and ethical considerations which hold back the broader practical use of this technology. In this review, we describe some of the epigenetic features of the early mammalian embryo, which might shed light on the mechanisms that enable and regulate totipotency.

Totipotency and pluripotency

Totipotency can be defined on the basis of cellular potency. A plausible definition for totipotency is the ability of a cell to contribute to all lineages in an organism. This contrasts to pluripotency, which is the potential to differentiate into all three germ layers of the embryo, but not all lineages. According to the criteria

defined above, blastomeres up to the 4-cell and 8-cell stage are totipotent in the mouse. This was demonstrated by transplantation experiments in which single blastomeres were shown to be able to contribute to all lineages in the embryo when supported by carrier blastomeres (3-5). However, it seems that not all blastomeres at the 4-cell or 8-cell stage display complete totipotency because some bias of a single blastomere towards contributing to a specific lineage has been observed at these stages (6-8). Plachta and colleagues have documented an intraembryonic variation that emerges after the 4-cell stage in the molecular kinetics of OCT4. The slower kinetics of OCT4, which were interpreted as tighter binding to chromatin, were found to correlate with specification of the ICM (9). Indeed, the possibility remains that when transplanted, a blastomere is 'reprogrammed' to regain totipotency because it is surrounded by carrier cells. Thus, totipotency can be defined more stringently as the ability of a single cell to develop into a complete organism by itself. Therefore in the mouse only the zygote and 2-cell stage blastomeres have been clearly shown to fulfil such stringent criteria (10). It is important to consider the similarities and differences between pluripotent and totipotent cells. Pluripotent cells can self-renew, i.e., ES and iPS cells can be stably maintained in culture. In addition, a pluripotent state can be induced experimentally (1, 11, 12). Therefore, future research directions should address whether totipotency can be induced *in vitro*. And if so, what are the signalling features or culture requirements for such? The potential of such a possibility is enormous, as this will certainly set the framework to underscore mechanisms underlying larger plasticity and therefore expand our reach for therapeutical approaches. Indeed, we consider that the larger plasticity inherent to totipotent cells will make generation of different cell types a much more efficient process once we uncover the mechanisms that regulate transit between different cellular states. Herein, we review the unique features of the early mouse embryo, which might shed light on the basis of cell totipotency.

(De)marking DNA (de)methylation

Early mouse embryos are characterised by a unique chromatin state that underlies epigenetic reprogramming, a process which is triggered upon fertilisation and is believed to be essential for the reacquisition of totipotency. Recent analysis of DNA methylation in early embryo has provided a more thorough understanding of that process. Just after fertilisation, the paternal genome is subject to a degree of active demethylation and hydroxymethylation, while the maternal genome is subject to slower passive demethylation (13-15). Consistently, when methylated DNA is visualised by immunostaining, the maternal genome displays a higher methylation level compared to the paternal genome in the zygote (14). However, quantitation of methylation levels is not accurate through immunostaining, and this approach does not provide locus-specific methylation insights. Recent DNA methylation mapping at a base pair resolution tackled these problems (16, 17). The results indicate that the oocyte genome is kept in a rather hypomethylated state and that only subtle global changes in DNA methylation from the zygote to

the 8-cell stage occur, apart from a significant reduction of paternal DNA methylation during the transition between the sperm and zygote. Thus, preimplantation embryos are characterised by global DNA hypomethylation. This finding might be a clue to understand how a totipotent state is generated and perhaps can give some hints into how to induce a totipotent cell by manipulating DNA methylation at specific loci, primarily those loci that become most drastically demethylated such as retrotransposons.

Asymmetric chromatin organisation

Because the sperm and oocyte chromatin are organised under different spatial constraints before fertilisation, the two parental genomes undergo different processes of chromatin reorganisation after fertilisation, which are accompanied by asymmetric histone modification patterns. For example, H3K9me3, a constitutive heterochromatin mark, is more prominent in the maternal pronucleus (18, 19). This asymmetric pattern of histone modification is maintained at least until the 8-cell stage (20). Furthermore, the asymmetry in histone modifications is mechanistically related to the asymmetry in DNA methylation and demethylation. Indeed, Nakamura et al. demonstrated recently that specific association of Stella with H3K9me2 on the maternal chromatin protects the maternal genome from Tet-mediated hydroxymethylation, providing a mechanistic explanation as to why the paternal genome only is subject to a drastic Tetdependent hydroxymethylation (21). Because this asymmetric chromatin organisation characterises the totipotent cells in the early embryo, the question that follows is whether this asymmetric organisation can influence totipotency. Although this feature is uniquely found in the early embryo, it seems that the asymmetry per se might not be involved in totipotency because totipotent cells generated by SCNT have a single nucleus (and therefore no obvious parental asymmetry) (22). Thus, although the parental epigenetic asymmetry as such may not be required for totipotency, we argue that a special epigenetic status that regulates cell potency must exist in totipotent cells. In particular results generated in the recent years have indicated that the early embryo has not only a unique chromatin organisation but also a high transcriptional activity of heterochromatic regions, which might contribute to generating an epigenetic environment permissive for totipotency.

Chromatin organisation: nucleolus precursor bodies and chromocentre formation

In somatic cells, centromeric and pericentromeric chromatin form the chromocentres, which are typically identified by intense signals of dyes such as DAPI and Hoechst, in mouse and to a lesser extent in human cells. However, the embryo starts its development without organising pericentromeric heterochromatin into chromocentres. Instead, the early embryo develops a unique nuclear structure in which unusual structures of compact fibrillar spheres, the Nucleolar Precursor Bodies (NPBs), are formed in both pronuclei after fertilisation and

persist until at least the 4-cell stage (23-26). At subsequent stages, the chromocentres form concomitantly with the disappearance of NPBs (26, 27). NPB-like bodies also form in SCNT embryos, which is an indicator of cloning efficiency and reprogramming, suggesting that the ooplasm has the ability to generate this structure (24, 28). This nucleolar structure is inherited from the oocyte but not from the sperm, and development arrests at 2-cell or 4-cell stage when NPBs are physically removed (29). Such a chromatin organisation around the nucleolus in the primary oocyte is predictive of their developmental competence: oocytes with a nucleolus around which heterochromatin forms a ring have higher potential to develop beyond the 2-cell stage (30-32), suggesting that the special nuclear structure and chromatin organisation provided maternally is essential for development and cell potency. NPBs are associated with centromeric and pericentromeric chromatin and as such are decorated by some heterochromatin marks like H3K9me3, H3K27me3, H3K64me3 and H4K20me3 in a parent-of-origin and stage-specific manner, suggesting that NPBs might be involved in heterochromatin remodelling or formation (19, 20, 24, 26, 33-35]. Surprisingly however, a recent work showed that an active chromatin mark, H3K36me3, is also enriched in these regions in the female pronucleus (36). The pericentromeric chromatin seems to be indeed subject to dynamic changes. In particular, the major satellites of pericentromeric chromatin are transcribed in the zygote and 2-cell stage embryos in a strand-specific manner, but become subsequently silenced, suggesting that the epigenetic status of major satellite repeats is dynamically controlled during early development (20, 25, 34). Furthermore, there is a clear temporal correlation between the reorganisation of NPBs into chromocentres and the transition towards pluripotency, but it is unclear whether such global changes in nuclear architecture are functionally linked with changes in developmental plasticity. Some indirect indication that this might be the case comes from recent analysis of the localisation of telomeres and centromeres throughout development, which reveals that reprogramming includes remarkable changes of chromosome organisation at the 4-cell stage (27). Indeed, we believe that the 4-cell stage constitutes an important 'transitional' state in terms of features of the embryonic chromatin, which might be relevant for subsequent cell fate specification and differentiation. In the future, it will be intriguing to examine the relevance of NPBs in establishing and maintaining cell totipotency after fertilisation. For example, identifying factors responsible for their formation and/or their structural components might shed some light in the molecular mechanisms underlying this peculiar embryonic chromatin organisation. The chaperone nucleoplasmin2 (NPM2) is one of the components of NPBs, but very little is known about other factors that might be associated to the NPBs. Deletion of *Npm2* in the oocyte results in embryos with abnormal nuclear morphology lacking NPBs and in arrested development, suggesting that NPBs impact global nuclear organisation and that this organisation is essential for embryonic development (37). Furthermore, it will be interesting to test if components of the NPBs are involved in physically targeting heterochromatic chromatin to the NPBs and whether manipulating such targeting might affect cell plasticity.

(Re-)activation of retrotransposons

As mentioned above, histone modifications characteristic of constitutive heterochromatin are detected only in the maternal pronucleus (19, 33, 38). These constitutive heterochromatin marks are maintained at a low level, especially on the paternal chromatin, at least until the 8-cell stage (20, 35). Probably because of this unique chromatin landscape, in addition to the hypomethylated DNA status, some retrotransposon elements are highly transcribed at the earliest developmental stages and in contrast to most cell types, where these elements are silenced. Retrotransposons are broadly classified into either long terminal repeat (LTR)-containing (including endogenous retroviruses, ERVs) or non-LTR retrotransposons such as LINE and SINE (long and short interspersed nuclear elements, respectively). The transcriptional kinetics of specific retrotransposon elements varies during early development. Line L1s and IAPs (a family of LINE superfamily and a subfamily of LTR-containing ERV-K family, respectively) are transcriptionally activated after fertilisation, and their expression seems to be maintained during preimplantation development (39-41). MaLR and MERV-L transcripts, both of them LTR-retrotransposons, are specifically abundant in mature oocytes and 2-cell stage embryos, respectively (39, 40, 42-44). Recent DNA methylation analysis shows that specific Line L1 and LTR retrotransposon subfamilies are extensively demethylated during preimplantation development, but IAP elements are resistant to demethylation (17). This implies that retrotransposon expression in the embryo cannot be explained solely by their DNA methylation status. Indeed, in mouse ESCs (mESCs), although depletion of G9a, a histone H3K9me2 methyltransferase, causes dramatic loss of DNA methylation at retrotransposons without affecting H3K9me3, expression of retrotransposons is only slightly affected (45-47). Conversely, depletion of Eset/Setdb1, another H3K9 methyltransferase, in mESCs does not affect DNA methylation at retrotransposons globally while H3K9me3 marks are lost, resulting in the derepression of class I and II ERVs (46, 48). All these data suggest that the expression of different retrotransposon classes is controlled by slightly different mechanisms. Furthermore, the fact that their dynamics of expression varies among retrotransposon families or subfamilies during early development opens up the possibility of the existence of yet-to-be discovered sequence specific factor(s) that recognise individual retrotransposons to silence or activate them. Indeed, a transcription factor-based model for the silencing of the repetitive sequences of the major satellites was recently put forward, and some retroelements function as enhancers recognised by transcription factors (49, 50). Potentially, such transcription factors might be responsible for the transcriptional regulation of single copy genes as well as retroelements, because derepression of retroelements induces the expression of 'chimeric' transcripts in embryos (43, 51). Such factors for silencing could include KRAB-ZFPs, of which a few hundred are encoded in the mouse genome and can both, target specific sequences and recruit histone methyltransferase and histone deacetylase through KAP1 (52-54). Importantly, recent work demonstrated that the maternal load of KAP1 plays a key role in preserving epigenetic stability in the pre-implantation embryo

(55). Thus, the specific transcriptional landscape of retrotransposons in the early embryo might have a direct regulatory effect on totipotency. It will therefore be important to establish whether the activation of retrotransposons that occurs after fertilisation is a side effect of reprogramming or whether it is a necessary step in the reacquisition of totipotency, or even more, whether activation of specific retrotransposons could be the trigger to establish totipotency.

Identification of a '2-cell-like' stem cell

Totipotent and pluripotent cells have the characteristic feature of retrotransposon expression. Recently, a remarkable discovery was reported, whereby a small percentage (0.5%) of mES cells display a transcriptional profile that is similar to that of the 2-cell stage embryo, and they have low levels of Sox2 and undetectable Oct4 protein. When this '2-cell-like' ES cells were injected into blastocysts, they contributed to both embryonic and extraembryonic tissues (56). This is not a property of mES cells, which are unable to contribute to all lineages unless they are genetically manipulated. In that study, the expression of the retrotransposon MERV-L, which is highly expressed in 2-cell stage embryos, was used to identify 2-cell-like cells in mESCs cultures (56). These cells have high level of active histone marks like H3K4me2 and represent a rare transient population of mESCs. It therefore remains unclear whether these totipotent-like cells can be stably maintained in culture. However, treatment with an HDAC inhibitor or the absence of the silencing histone modifiers Kdm1a, Kap1 or G9a, facilitates the emergence of 2-cell like cells marked by MERV-L expression. Indeed, Kap1 depletion leads to upregulation of MERV-L (57). This finding raises the interesting possibility that 2-cell like cells can be induced and/or maintained through modifying the epigenetic state and/or expression of retrotransposons. This is however, unlikely to be a straightforward inference, since knockdown of Eset/Setdb1 in mESCs or 2-cell stage embryos induces cells to differentiate into the extraembryonic lineage (46, 58, 59). Thus, in this case, Eset/Setdb1 depletion results in the loss of pluripotency. In conclusion, it would seem that artificially inducing an open chromatin state facilitates cells towards the totipotent state but a finely tuned or a metastable chromatin state is necessary for totipotency (60). Along these lines, some important questions that follow are how 2-cell like cells arise in culture, whether this is a regulated process or more stochastic and if their state can be stably maintained in culture. Interestingly, a recent report documents that cells capable to form trophoblast and ICM, referred as totipotent, arise in ES cell cultures grown in 2i medium (61). These cells are identified on the basis of the expression of the early endoderm differentiation marker Hex and unlike the ERV-marked ES cells, they express pluripotency and differentiation factors simultaneously and therefore look more similar to those in 16-cell/early morula. Considering the two definitions of totipotency we discussed above, it will be important to address if any difference in the cell potency between the 2-cell like cells and these Hex+ cells exists. Could it be then that the expression of repetitive elements has some active role for totipotency? Although this has not been addressed thoroughly, there is some

evidence showing that microinjecting antisense oligonucleotides targeting Line L1s into zygotes causes developmental arrest at 2/4-cell stage (62). Alternatively, the expression of repetitive elements might be a side-effect consequence of the unique epigenetic state and the gene expression pattern for totipotency or pluripotency that is shaped at the expense of the risk of retrotransposition. Further understanding of the chromatin state during early development will provide important clues to decipher the mechanism for generating and maintaining totipotency and a potential role for the expression of retroelements in regulating these states. Moreover, whether this transient nature of the 2-cell like cells in mES culture is also a feature of human ES cells will be an important question to address.

Concluding remarks and future research directions

Cell plasticity is a fundamental pillar for developmental and stem cell biology. Understanding the molecular mechanisms behind maintenance and manipulation of cell fate has become an intense research area in the past years. Key advances including the generation of iPS cells and SCNT experiments have provided us with tools and models to explore the underlying molecular mechanisms of cell potency. The use of different model systems should continue to reveal common strategies that have evolved to regulate cell plasticity. For example, the histone variant H3.3 has been recently shown to be required for efficient SCNT (63) and has been also ascribed a role in reprogramming chromatin signatures after fertilisation in the mouse (34). Addressing the molecular mechanisms of cell plasticity and cell fate *in vivo* during development must remain an important research focus, in particular, since it will unveil the pathways that nature has selected to generate cells with an extremely high degree of plasticity. The development of quantitative approaches to map the epigenome (64) and to profile quantitatively single cells (65, 66) is a very important step towards that goal. Here, we have attempted to define a framework of the molecular features of totipotency based on recent findings, which in our view is essential to understand how pluripotent cells form *in vivo*. Indeed, losing totipotency is a necessary step to form pluripotent cells.

Bibliografia

1. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006; 126: 663-676.
2. Tachibana M, Amato P, Sparman M, Gutierrez NM, Tippner-Hedges R, Ma H, Kang E, Fulati A, Lee HS, Sritanaudomchai H, et al. Human embryonic stem cells derived by somatic cell nuclear transfer. *Cell*. 2013; 153:1228-1238.
3. Balakier H, Pedersen RA. Allocation of cells to inner cell mass and trophectoderm lineages in preimplantation mouse embryos. *Dev Biol*. 1982; 90: 352-362.
4. Kelly SJ: Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres. *J Exp Zool*. 1977; 200:365-376.

5. Rossant J. Postimplantation development of blastomeres isolated from 4- and 8-cell mouse eggs. *J Embryol Exp Morphol.* 1976; 36: 283-290.
6. Fujimori T, Kurotaki Y, Miyazaki J, Nabeshima Y. Analysis of cell lineage in two- and four-cell mouse embryos. *Development.* 2003; 130: 5113-5122.
7. Piotrowska-Nitsche K, Perea-Gomez A, Haraguchi S, Zernicka-Goetz M. Four-cell stage mouse blastomeres have different developmental properties. *Development.* 2005; 132: 479-490.
8. Tabansky I, Lenarcic A, Draft RW, Loulier K, Keskin DB, Rosains J, Rivera-Feliciano J, Lichtman JW, Livet J, Stern JN et al. Developmental bias in cleavage-stage mouse blastomeres. *Curr Biol.* 2013; 23: 21-31.
9. Plachta N, Bollenbach T, Pease S, Fraser SE, Pantazis P. Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. *Nat Cell Biol.* 2011; 13: 117-123.
10. Tarkowski AK. Experiments on the development of isolated blastomeres of mouse eggs. *Nature.* 1959; 184: 1286-1287.
11. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature.* 1981; 292: 154-156.
12. Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA.* 1981; 78: 7634-7638.
13. Iqbal K, Jin SG, Pfeifer GP, Szabo PE. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci USA.* 2011; 108: 3642-3647.
14. Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* 2002; 241: 172-182.
15. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W, Walter J. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun* 2011; 2:241.
16. Smallwood SA, Tomizawa S, Krueger F, Ruf N, Carli N, Segonds-Pichon A, Sato S, Hata K, Andrews SR, Kelsey G. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat Genet.* 2011; 43: 811-814.
17. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 2012, 484: 339-344. The authors apply reduced representation bisulphite sequencing and provide the first genome scale coverage of DNA methylation patterns with single base resolution revealing different demethylation phases of different genomic regions.
18. Arney KL, Bao S, Bannister AJ, Kouzarides T, Surani MA. Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int J Dev Biol.* 2002; 46: 317-320.
19. Santos F, Peters AH, Otte AP, Reik W, Dean W: Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev Biol.* 2005; 280: 225-236.
20. Puschendorf M, Terranova R, Boutsma E, Mao X, Isono K, Brykczynska U,

- Kolb C, Otte AP, Koseki H, Orkin SH et al. PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat Genet.* 2008; 40: 411-420.
21. Nakamura T, Liu YJ, Nakashima H, Umehara H, Inoue K, Matoba S, Tachibana M, Ogura A, Shinkai Y, Nakano T. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* 2012, 486:415-419. This work shows that PGC7/Stella binds preferentially histone H3 tails dimethylated at K9 and it is suggested as a mechanism that protects the maternal genome from demethylation/hydroxymethylation.
 22. Wakayama T, Perry AC, Zuccotti M, Johnson KR, Yanagimachi R. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature.* 1998; 394: 369-374.
 23. Flechon JE, Kopecny V. The nature of the 'nucleolus precursor body' in early preimplantation embryos: a review of finestructure cytochemical, immunocytochemical and autoradiographic data related to nucleolar function. *Zygote.* 1998; 6: 183-191.
 24. Martin C, Beaujean N, Brochard V, Audouard C, Zink D, Debey P. Genome restructuring in mouse embryos during reprogramming and early development. *Dev Biol.* 2006; 292: 317-332.
 25. Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, Almouzni G. A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev Cell.* 2010; 19: 625-638.
 26. Probst AV, Santos F, Reik W, Almouzni G, Dean W. Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma.* 2007; 116: 403-415.
 27. Aguirre-Lavin T, Adenot P, Bonnet-Garnier A, Lehmann G, Fleurot R, Boulesteix C, Debey P, Beaujean N: 3D-FISH analysis of embryonic nuclei in mouse highlights several abrupt changes of nuclear organization during pre-implantation development. *BMC Dev Biol.* 2012; 12: 30. This paper includes a remarkably detailed and controlled analysis of the topology of heterochromatic regions in the early mouse embryo and suggests global rearrangement of the embryonic genome at the 4-cell stage.
 28. Maalouf WE, Liu Z, Brochard V, Renard JP, Debey P, Beaujean N, Zink D. Trichostatin A treatment of cloned mouse embryos improves constitutive heterochromatin remodeling as well as developmental potential to term. *BMC Dev Biol.* 2009; 9: 11.
 29. Ogushi S, Palmieri C, Fulka H, Saitou M, Miyano T, Fulka J Jr: The maternal nucleolus is essential for early embryonic development in mammals. *Science.* 2008; 319: 613-616.
 30. Liu H, Aoki F. Transcriptional activity associated with meiotic competence in fully grown mouse GV oocytes. *Zygote.* 2002; 10: 327-332.
 31. Zuccotti M, Giorgi Rossi P, Martinez A, Garagna S, Forabosco A, Redi CA. Meiotic and developmental competence of mouse antral oocytes. *Biol Reprod.* 1998; 58: 700-704.

32. Zuccotti M, Ponce RH, Boiani M, Guizzardi S, Govoni P, Scandroglio R, Gargana S, Redi CA. The analysis of chromatin organisation allows selection of mouse antral oocytes competent for development to blastocyst. *Zygote*. 2002; 10: 73-78.
33. Daujat S, Weiss T, Mohn F, Lange UC, Ziegler-Birling C, Zeissler U, Lappe M, Schubeler D, Torres-Padilla ME, Schneider R. H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. *Nat Struct Mol Biol*. 2009; 16: 777-781.
34. Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ, Torres-Padilla ME. Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat Cell Biol*. 2010; 12: 853-862.
35. Wongtawan T, Taylor JE, Lawson KA, Wilmot I, Pennings S. Histone H4K20me3 and HP1alpha are late heterochromatin markers in development, but present in undifferentiated embryonic stem cells. *J Cell Sci*. 2011; 124: 1878-1890.
36. Boskovic A, Bender A, Gall L, Ziegler-Birling C, Beaujean N, Torres-Padilla ME. Analysis of active chromatin modifications in early mammalian embryos reveals uncoupling of H2A.Z acetylation and H3K36 trimethylation from embryonic genome activation. *Epigenetics*. 2012; 7: 747-757.
37. Burns KH, Viveiros MM, Ren Y, Wang P, DeMayo FJ, Frail DE, Eppig JJ, Matzuk MM. Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos. *Science*. 2003; 300: 633-636.
38. Kourmouli N, Jeppesen P, Mahadevhaiah S, Burgoyne P, Wu R, Gilbert DM, Bongiorno S, Prantera G, Fanti L, Pimpinelli S et al. Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. *J Cell Sci*. 2004; 117: 2491-2501.
39. Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, Carninci P, Torres-Padilla ME. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat Struct Mol Biol*. 2013; 20: 332-338.
40. Inoue A, Matoba S, Zhang Y. Transcriptional activation of transposable elements in mouse zygotes is independent of Tet3-mediated 5-methylcytosine oxidation. *Cell Res*. 2012; 22: 1640-1649.
41. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, Kazazian HH Jr. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev*. 2009; 23: 1303-1312.
42. Kigami D, Minami N, Takayama H, Imai H. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biol Reprod*. 2003; 68: 651-654.
43. Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D, Knowles BB. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell*. 2004; 7: 597-606.
44. Svoboda P, Stein P, Anger M, Bernstein E, Hannon GJ, Schultz RM. RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos. *Dev Biol*. 2004; 269: 276-285.
45. Dong KB, Maksakova IA, Mohn F, Leung D, Appanah R, Lee S, Yang HW,

- Lam LL, Mager DL, Schubeler D et al. DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J.* 2008; 27: 2691-2701.
46. Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, Kimura H, Tachibana M, Lorincz MC, Shinkai Y. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature.* 2010; 464: 927-931.
47. Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y. G9a/ GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J.* 2008; 27: 2681-2690.
48. Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, Shinkai Y, Mager DL, Jones S, Hirst M et al. DNA methylation and SETDB1/ H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell.* 2011; 8: 676-687.
49. Bulut-Karslioglu A, Perrera V, Scaranaro M, de la Rosa-Velazquez IA, van de Nobelen S, Shukeir N, Popow J, Gerle B, Opravil S, Pagani M et al. A transcription factor-based mechanism for mouse heterochromatin formation. *Nat Struct Mol Biol.* 2012; 19: 1023-1030.
50. Chuong EB, Rumi MA, Soares MJ, Baker JC. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat Genet.* 2013; 45: 325-329. This study demonstrates that a species-specific retrotransposon has enhancer activity in specific tissue, implying that retroelements might have shaped the regulatory elements of the genome during evolution. This is an excellent example of the regulatory function of retroelements which account for a large portion of our genome.
51. van de Lagemaat LN, Landry JR, Mager DL, Medstrand P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* 2003; 19: 530-536.
52. Emerson RO, Thomas JH. Adaptive evolution in zinc finger transcription factors. *PLoS Genet.* 2009; 5: e1000325.
53. Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ III. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* 2002; 16: 919-932.
54. Schultz DC, Friedman JR, Rauscher FJ III. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. *Genes Dev.* 2001; 15: 428-443.
55. Messerschmidt DM, de Vries W, Ito M, Solter D, Ferguson-Smith A, Knowles BB. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science.* 2012; 335: 1499-1502. The deletion of Trim28/Kap1 in the oocyte results in a highly variable phenotype resulting at least partially from loss of imprinting in the early embryo. This study provides a molecular basis of a clear epigenetic phenotype.
56. Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. Embryonic stem cell potency fluctuates

- with endogenous retrovirus activity. *Nature*. 2012; 487: 57-63. The authors show that a rare transient population of mESCs has a potential to contribute to all lineages. These cells are identified by MERV-L retrotransposon expression which is also induced at 2-cell stage *in vivo*, and have a global RNA profile similar to 2-cell embryos. This study opens up the potential to maintain and/or induce totipotent cells *in vitro*.
57. Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T, Mailard PV, Layard-Liesching H, Verp S, Marquis J et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature*. 2010; 463: 237-240.
 58. Bilodeau S, Kagey MH, Frampton GM, Rahl PB, Young RA. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev*. 2009; 23: 2484-2489.
 59. Yuan P, Han J, Guo G, Orlov YL, Huss M, Loh YH, Yaw LP, Robson P, Lim B, Ng HH. Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev*. 2009; 23: 2507-2520.
 60. Furusawa C, Kaneko K: A dynamical-systems view of stem cell biology. *Science*. 2012; 338: 215-217.
 61. Morgani SM, Canham MA, Nichols J, Sharov AA, Migueles RP, Ko MS, Brickman JM. Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell Rep*. 2013; 3: 1945-1957.
 62. Beraldi R, Pittoggi C, Sciamanna I, Mattei E, Spadafora C: Expression of LINE-1 retroposons is essential for murine preimplantation development. *Mol Reprod Dev* 2006; 73: 279- 287.
 63. Jullien J, Astrand C, Szenker E, Garrett N, Almouzni G, Gurdon JB. HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to *Xenopus* oocytes. *Epigenetics Chromatin* 2012, 5:17. This study shows that the histone variant H3.3 provided by the oocyte is strictly required for reprogramming upon SCNT. The other H3 variant H3.2, is dispensable.
 64. Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, Balasubramanian S. Quantitative sequencing of 5- methylcytosine and 5- ydroxymethylcytosine at single-base resolution. *Science* 2012; 336: 934-937.
 65. Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell* 2010; 18: 675-685.
 66. Lorthongpanich C, Doris TP, Limviphuvadh V, Knowles BB, Solter D. Developmental fate and lineage commitment of singled mouse blastomeres. *Development* 2012; 139: 3722-3731.

La natura epigenetica dei centromeri di mammifero

Elena Raimondi

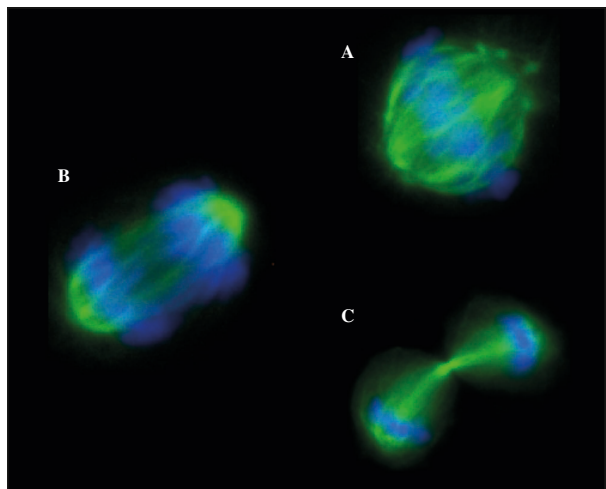
Dipartimento di Biologia e Biotecnologie, Università di Pavia

La fedele trasmissione del materiale ereditario nelle cellule figlie ad ogni divisione cellulare è un requisito fondamentale per la vitalità di cellule e organismi. Il DNA eucariotico è associato alle proteine istoniche e super-avvolto a formare i cromosomi, che devono essere replicati e segregare correttamente durante la divisione cellulare. La regione specializzata, che garantisce la corretta segregazione del materiale genetico nelle cellule eucariotiche, è il centromero. Il centromero recluta una macro-struttura proteica, chiamata cinetocoro, che ne media l'interazione con il fuso mitotico e meiotico.

Da un punto di vista citologico, il centromero della maggior parte dei vertebrati appare come una costrizione primaria, ben visibile durante la divisione cellulare, che mantiene uniti i cromatidi fratelli fino alla metafase e ne media la migrazione ai poli durante l'anafase (Figura 1).

Anche se il centromero di molti organismi è stato descritto in modo dettagliato sia da un punto di vista citologico sia genetico, sorprendentemente, la natura molecolare della funzione centromerica resta un problema aperto e una delle più affascinanti sfide della genetica e biologia molecolare.

Fig. 1 - Divisione cellulare - In blu i cromosomi colorati con DAPI, in verde il fuso mitotico evidenziato con anticorpi anti beta tubulina. A) Metafase, B) anafase, C) telofase.



Il primo aspetto sorprendente è che, nonostante la funzione centromerica sia altamente conservata in tutti gli eucarioti, le sequenze di DNA centromerico sono altamente divergenti sia tra specie sia entro specie. È stato ampiamente dimostrato che non esiste una sequenza di DNA necessaria e sufficiente per sostenere la funzione centromerica. La contraddizione tra conservazione evolutiva della funzione centromerica e divergenza delle sequenze di DNA ad essa sottostanti è nota come “paradosso del centromero” (1, 2). Le prime prove dirette del fatto che la sequenza di DNA non determina la funzione centromerica derivano dall’analisi di anomalie cromosomiche. L’esistenza di cromosomi umani dicentrici, con centromeri identici quanto a sequenza, uno solo dei quali è funzionale, dimostra che il DNA centromerico, da solo, non è sufficiente per specificare la funzione (3). Di contro, esistono centromeri privi delle tipiche sequenze di DNA ripetuto centromerico, eppure completamente funzionali, a dimostrazione del fatto che particolari sequenze di DNA non sono necessarie per determinare la funzione (4).

La caratteristica unificante di tutti i centromeri eucariotici è la presenza, a livello della cromatina centromerica, di un istone H3 modificato, la proteina CENP-A o CENH3 (Figura 2); inoltre, diverse osservazioni sperimentali dimostrano che l’instaurarsi della funzione centromerica richiede cambiamenti conformazionali della cromatina centromero-specifici ed è pertanto di natura prettamente epigenetica (5).

Per quanto, la funzione centromerica non dipenda dalla sequenza di DNA sottostante, i centromeri della maggior parte dei metazoi sono associati a estesi array di sequenze di DNA altamente ripetuto (DNA satellite) che possono estendersi per diversi milioni di coppie di basi. La presenza di sequenze di DNA altamente ripetuto a livello dei centromeri ne ha sempre ostacolato la dettagliata analisi molecolare ed il sequenziamento completo. In questo scenario, ha avuto importanza fondamentale la scoperta, da parte del nostro gruppo di ricerca, che il centromero di un cromosoma di cavallo, il cromosoma 11, è completamente privo di DNA

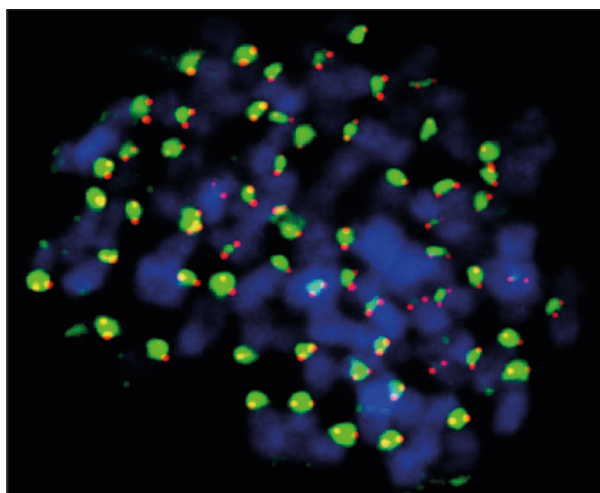


Fig. 2 - La proteina CENP-A-
In blu i cromosomi colorati con DAPI; in verde l’eterocromatina centromerica; in rosso la proteina CENP-A evidenziata con un anticorpo monoclonale.

satellite (6). I nostri dati indicano che il centromero del cromosoma 11 di cavallo è evolutivamente recente e rappresenta una fase precoce nel processo di maturazione del centromero. La scoperta successiva, nelle specie appartenenti al genere *Equus*, di un numero eccezionalmente elevato di centromeri privi di DNA satellite (7, 8), ne fa un modello biologico unico per lo studio della nascita, evoluzione e completa maturazione del centromero, dimostrando, in modo inequivocabile, che la funzione centromerica è indipendente dalla sequenza sottostante.

La componente epigenetica distintiva dei centromeri è, come accennato in precedenza, la proteina CENP-A. Poiché CENP-A è un istone modificato, i nucleosomi che la contengono sono conformazionalmente più rigidi. Nel “core” centromerico funzionale, nucleosomi contenenti CENP-A e nucleosomi contenenti l'istone H3 si alternano in modo non regolare (9). Per quanto la dimensione del dominio occupato da nucleosomi contenenti la proteina CENP-A sia variabile, rimane abbastanza costante il numero totale di nucleosomi CENP-A presenti al centromero.

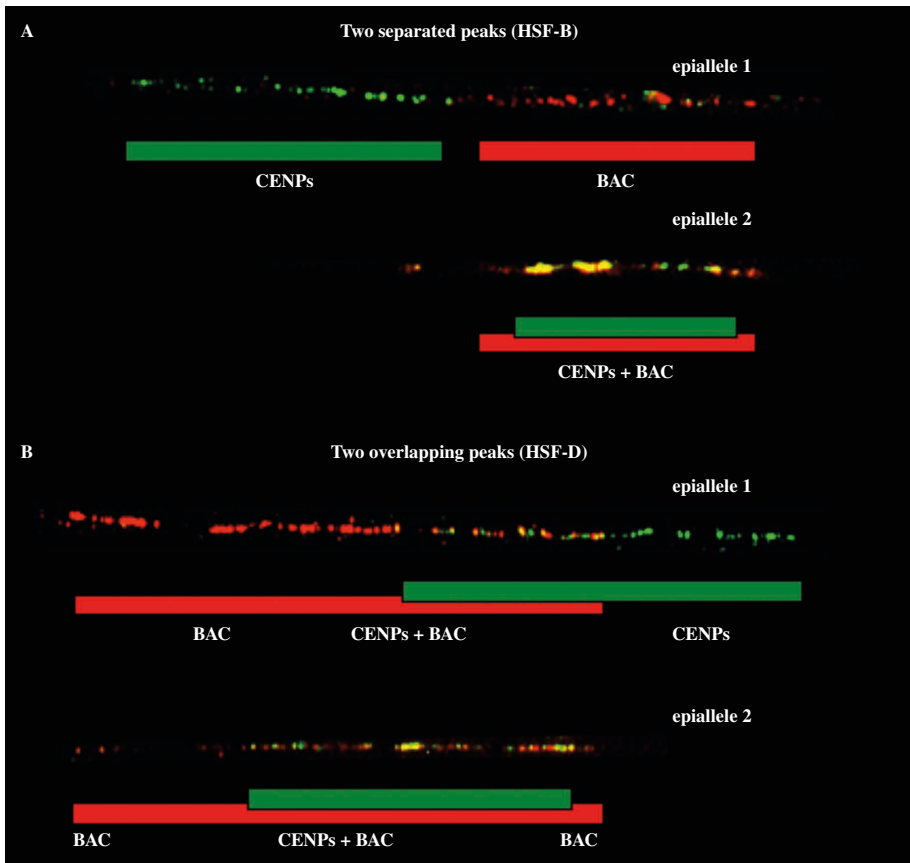


Fig. 3 - Esperimenti di immunoFISH su fibre di cromatina estesa per l'analisi del fenomeno dello “sliding centromerico” - In rosso i BAC usati come sonde di FISH. In verde la proteina CENP-A evidenziata con immunofluorescenza. A) alleli con domini funzionali separati; B) alleli con domini parzialmente sovrapposti.

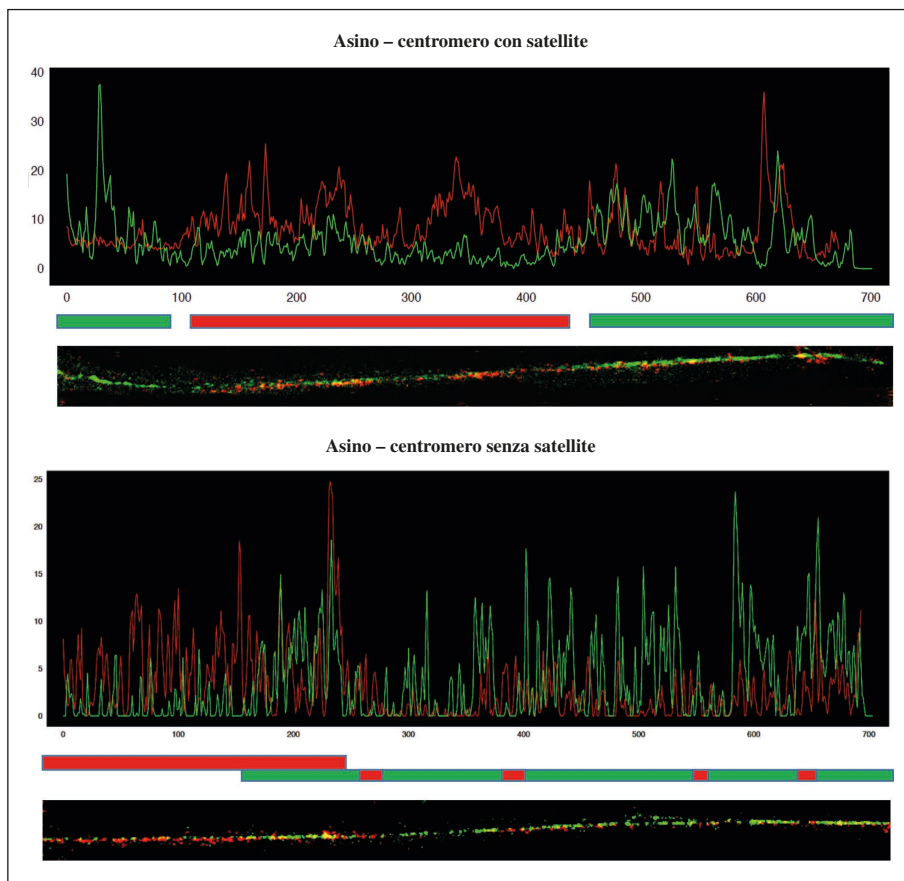


Fig. 4 - Modificazioni epigenetiche in centromeri di asino con DNA satellite (a) e senza DNA satellite (b) - In alto in entrambi i pannelli è rappresentata la quantificazione dei segnali fluorescenti. In basso in ogni pannello sono mostrate fibre di DNA esteso analizzate con esperimenti di doppia immuno-fluorescenza. In rosso è marcata la proteina CENP-A, in verde la modificazione istonica H3K9me3.

Queste osservazioni portano a pensare che esista un “valore soglia” di nucleosomi CENP-A compatibile con l’instaurarsi della funzione centromerica (10).

La cromatina centromerica è sempre stata ritenuta altamente condensata e trascrizionalmente inerte: un esempio tipico di eterocromatina costitutiva. Al contrario, è oggi opinione unanime che la “centrocromatina” si trovi in una conformazione all’interfaccia tra quella tipica dell’eterocromatina e quella tipica dell’euromatina (11). In altre parole, affinché un *locus* possa assumere funzione centromerica è necessario un equilibrio centromero-specifico tra modificazioni istoniche post-traduzionali caratteristiche della cromatina trascrizionalmente inerte e modificazioni tipiche della cromatina trascrizionalmente competente. Marcatori tipici della forma aperta della cromatina, come la dimetilazione della lisina 36 dell’istone H3 (H3K36me2) (12), sono presenti nel core centromerico; di contro, mancano altre modificazioni tipicamente euromatiche come l’aceti-

lazione degli stoni H3 e H4 e non sono neppure presenti modificazioni tipiche dell'eterocromatina costitutiva come la trimetilazione della lisina 9 dell'istone H3 (H3K9me3) (13).

Gli studi più recenti hanno dimostrato che la competenza trascrizionale è una caratteristica distintiva della cromatina centromerica indicando che questa prerogativa sia prerequisito per il reclutamento della proteina CENP-A (14). Trascritti centromerici e pericentromerici sono stati ritrovati in tutte le specie analizzate, dal lievito all'uomo. Si tratta di RNA non tradotti, che possono agire sia in cis sia in trans, il cui ruolo nella funzione centromerica non è del tutto chiaro. Nei mammiferi, sono stati trovati diversi trascritti centromerici di dimensione e con dinamica variabili. Nell'uomo, è stato identificato un "long non coding RNA", trascritto dal satellite centromerico, coinvolto nel reclutamento di CENP-A mediato dalla proteina HJURP (15), inoltre, è stato osservato che trascritti centromerici sono complessati con le proteine CENP-A e CENP-C (16).

Avvalendoci del modello biologico rappresentato dalle specie del genere *Equus*, abbiamo analizzato la natura dei centromeri privi di DNA satellite e confrontato le caratteristiche epigenetiche di questi con quelle dei centromeri contenenti le canoniche sequenze di DNA centromerico. I nostri studi hanno evidenziato che i centromeri privi di DNA satellite sono posizionalmente instabili grazie ad un fenomeno che abbiamo definito "scivolamento del centromero". In individui diversi e anche a livello dei cromosomi omologhi di un singolo individuo, il dominio di legame della proteina CENP-A può cambiare posizione generando "epialleli" ossia alleli funzionali. La funzione centromerica può instaurarsi in sequenze "centromerizzabili" entro una regione di circa 600 kb. (Figura 3) (17, 18).

Inoltre, l'analisi delle modificazioni istoniche in centromeri privi di DNA satellite e in centromeri con DNA satellite, ha dimostrato che i centromeri senza DNA satellite hanno modificazioni istoniche simili a quelle osservate a livello dei centromeri con DNA satellite anche se questi ultimi mostrano un'evidente riduzione di modificazioni tipicamente eterocromatiche (dati non pubblicati) (Figura 4). I nostri risultati preliminari suggeriscono che, al contrario di quelli contenenti DNA satellite, i centromeri che ne sono privi sono trascrizionalmente inerti. Un'ipotesi intrigante suggerisce che RNA trascritti dai centromeri con DNA satellite possano agire in trans complementando l'assenza di trascrizione dei centromeri senza satellite.

Bibliografia

1. Eichler EE. Repetitive conundrums of centromere structure and function. *Hum Mol Genet.* 1999; 8: 151-155.
2. Henikoff S, Ahmad K, Malik HS. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science.* 2001; 293: 1098-1102.
3. Earnshaw WC, Migeon BR. Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. *Chromosoma.* 1985; 92: 290-296.
4. Voullaire LE, Slater HR, Petrovic V, Choo KH. A functional marker cen-

- tromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? *Am J Hum Genet.* 1993; 52: 1153-1163.
5. McKinley KL, Cheeseman IM. The molecular basis for centromere identity and function. *Nat Rev Mol Cell Biol.* 2016; 17: 16-29.
 6. Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, Lear TL, Adelson DL, Bailey E, Bellone RR, et al. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science.* 2009; 326: 865-867.
 7. Piras FM, Nergadze SG, Poletto V, Cerutti F, Ryder OA, Leeb T, Raimondi E, Giulotto E. Phylogeny of horse chromosome 5q in the genus *Equus* and centromere repositioning. *Cytogenet Genome Res.* 2009; 126: 165-172.
 8. Piras FM, Nergadze SG, Magnani E, Bertoni L, Attolini C, Khoriauli L, Raimondi E, Giulotto E. Uncoupling of satellite DNA and centromeric function in the genus *Equus*. *PLoS Genet.* 2010; 6: e1000845.
 9. Smith L, Maddox PS. Reading the Centromere Epigenetic Mark. *Dev Cell.* 2017; 42: 110-112.
 10. Iwata-Otsubo A, Dawicki-McKenna JM, Akera T, Falk SJ, Chmátal L, Yang K, Sullivan BA, Schultz RM, Lampson MA, Black BE. Expanded Satellite Repeats Amplify a Discrete CENP-A Nucleosome Assembly Site on Chromosomes that Drive in Female Meiosis. *Curr Biol.* 2017; 27: 2365-2373.
 11. García Del Arco A, Erhardt S. Post-translational Modifications of Centromeric Chromatin. *Prog Mol Subcell Biol.* 2017; 56: 213-231.
 12. Bergmann JH, guez MGOMRI, Martins NMC, Kimura H, Kelly DA, Masumoto H, Larionov V, Jansen LET, Earnshaw WC. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO J.* 2011; 30: 328-340.
 13. Sullivan BA, Karpen GH. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol.* 2004; 11: 1076-1083.
 14. Perea-Resca C, Blower MD. Satellite Transcripts Locally Promote Centromere Formation. *Dev Cell.* 2017; 42: 201-202.
 15. Quenet, D., and Dalal, Y. A long non-coding RNA is required for targeting centromeric protein A to the human centromere. *Elife.* 2014; 3: e03254.
 16. McNulty SM, Sullivan LL, Sullivan BA. Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are Complexed with CENP-A and CENP-C. *Dev Cell.* 2017; 42: 226-240.e6.
 17. Purgato S, Belloni E, Piras FM, Zoli M, Badiale C, Cerutti F, Mazzagatti A, Perini G, Della Valle G, Nergadze SG, et al. Centromere sliding on a mammalian chromosome. *Chromosoma.* 2015; 124: 277-287.
 18. Nergadze SG, Piras FM, Gamba R, Corbo M, Cerutti F, McCarter JGW, Cappelletti E, Gozzo F, Harman RM, Antczak DF, Miller D, Scharfe M, Pavesi G, Raimondi E, Sullivan KF, Giulotto E. Birth, evolution and transmission of satellite-free mammalian centromeric domains. *Genome Research* 2018; accepted for publication.

Staminalità e patogenesi

Geppino Falco

Dipartimento di Biologia, Complesso Universitario Monte S. Angelo, Università di Napoli "Federico II"

Progenitor cells (PCs) are essential for the maintenance and regeneration of organs and tissues that exhibit high rates of cell turnover or regenerative reserve. Moreover, at the organism level, alterations in PCs homeostasis balance are clearly fundamental to physiological processes, such as organ development, and to pathological events, such as degenerative diseases, and cancer (1).

The widespread prevalence of debilitating diseases that result from pancreatic dysfunction is encouraging numerous studies to strengthen our knowledge of the principles governing the formation of this organ to promote the pancreatic cell regeneration, or to implement the cancer progression knowledge (2).

Many genes that have been shown to cause cancer were originally identified because of their role in embryonic development. The similarities between cancer and development are evident on many levels. Microscopically, cancerous tissues frequently appear as undifferentiated masses, with some tumor types exhibiting embryonic tissue organization. The increased mobility of cancer cells, leading to local invasion or metastasis, is reminiscent of the migratory behaviour of cells during development. At the molecular level, malignant tissues and developing

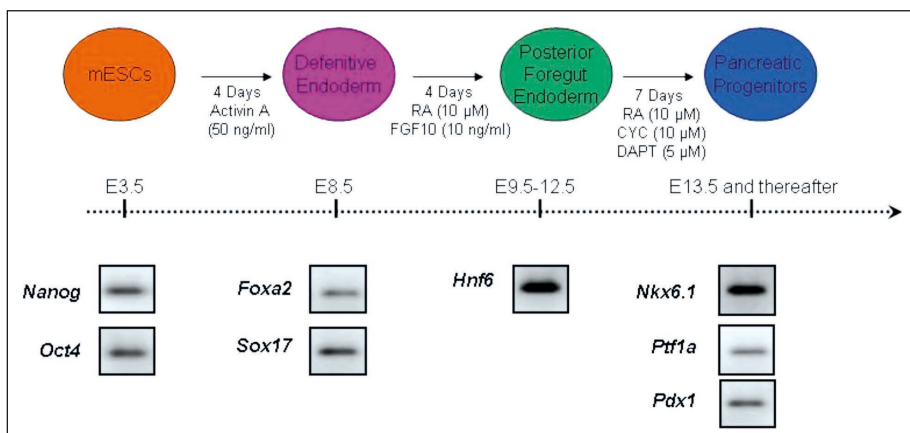


Fig. 1 - ESCs were cultured for 4 days treated with activin A, the definitive endoderm cells were cultured for another 4 days in the presence of both retinoid acid (RA) and FGF10. Following the posterior foregut endodermal stage, the cells were cultured in the presence of combined RA, cyclopamine (CYC), and N-N-(3,5-difluorophenacetyl)-L-alanylphenylglycine-butylester (DAPT) to give rise to pancreatic progenitors.

tissues have been shown to share signalling, transcription factor activity, regulation of chromatin structure, and DNA stability. One intriguing opportunity to challenge those issues comes from the studies of embryonic stem cells (ESCs) differentiation. ESCs are derived from the inner cell mass of blastocyst and are characterized by the remarkable peculiarity to produce the majority of cell types. Pluripotency confers the ability of ESCs to differentiate in ectoderm, endoderm, and mesoderm derivatives, producing the majority of cell types. Although the majority of ESCs divide without losing pluripotency, it has become evident that ESCs culture consists of multiple differentiating cell populations induced during spontaneous differentiation. ESCs expression heterogeneity reflects a stochastic transition without marking definitive commitment, such fluctuation represent a suitable *in vitro* model to improve the comprehension of the molecular mechanisms underlying early lineage fate (3).

The major aim of ESC-based protocols is to derive progenitor cells that retain both the capacity to proliferate and to differentiate following normal developmental ontogeny thus offering an unlimited source of safe precursor cells. The most promising differentiation protocols are based on extrinsic factors that are produced by the notochord, dorsal aorta, and pancreatic mesenchyme during pancreas organogenesis (Figure 1).

Advancements, such as ESCs direct pancreatic progenitor cell differentiation, efficient and rapid DNA genetic engineering, have greatly increased the ability to characterize biological dynamism in a tissue specific manner that it is likely to provide precious information on onset of pancreas pathology (Figure 2). The use of the cell line expressing reporter genes (such as GFP, beta-galactosidase etc)

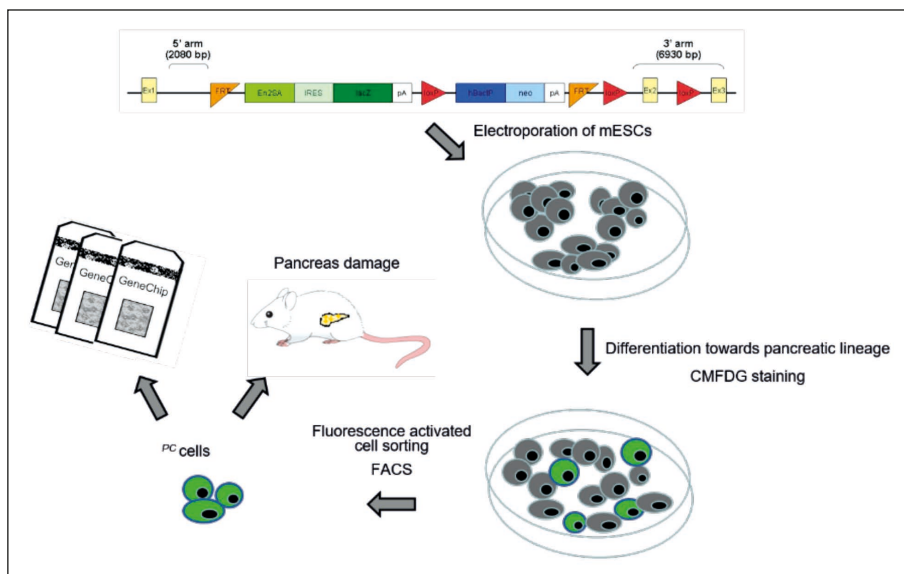


Fig. 2 - Scheme that shows how DNA editing in PC locus could enhance the visualization both *in vitro* and *in vivo* of regenerative homeostasis pathological condition.

under the promoter of the gene candidate offers the advantage to stain and collect the population homogeneously both *in vitro* and *ex vivo* (3).

In this way, nowadays it is possible to easily and efficiently characterize progenitor cell population. Importantly, the identification of a detailed molecular roadmap underlining the *in vivo* fate specification would further improve the efficiency and the quality of the *in vitro* cell derivatives. Actually, the characterization of those molecular factors is still missing because of the limited *in vivo* availability of progenitor cells (PPCs) (4).

In mouse, pancreatic organogenesis is a multistep progression made up of three stages: primary regulatory transition, from Embryonic Day 8.5 to Embryonic Day 10.5 (E8.5-E10.5), secondary regulatory transition (E13.5-E16.5) and third regulatory transition (after birth). A reliable source of PPCs is temporally and transiently enriched during pancreas budding processes, where PPCs are marked by multipotent phase occurring mainly during the *primary transition*. In the first developmental transition (E9.5-E12.5), the epithelium of the two buds under-

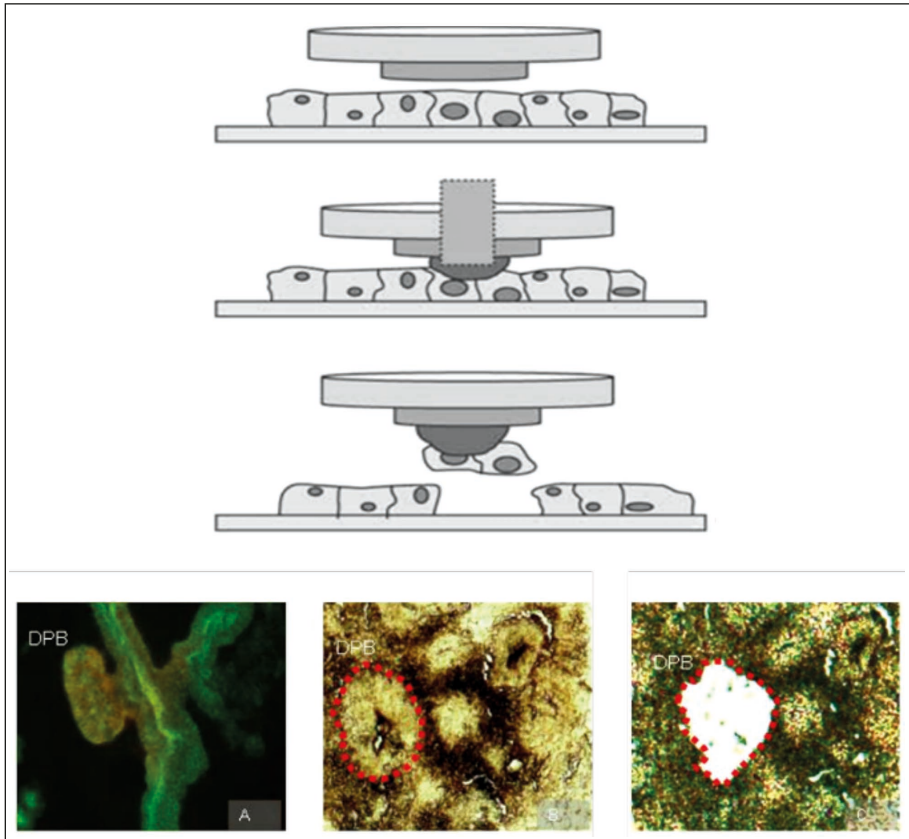


Fig. 3 - The top of the figure shows the rationale of LCM technology. The first panel shows the immunofluorescence of dorsal pancreas bud, the second panel shows section after dehydration and the third panel shows the sections after capture of dorsal pancreas bud.

goes branching morphogenesis, and the compartmentalization of the pancreatic epithelium commences, defining the epithelium into distinct “tip” and “trunk” domains. While the tip domains contain multipotential pancreatic cells (MPC), which are then believed to change into acinar-fated progenitors, the trunk domain consists of an endocrine-duct bipotential progenitor pool (5).

The *in vivo* identification of pancreas progenitor cell markers will be based on the analysis of gene expression profile of dorsal pancreatic buds isolated from E10.5 mouse embryos using Laser Capture Microdissection technique.

It is important to highlight that the molecular mechanisms that regulate the primary transition have been mainly focused on bud intrinsic signals but also extrinsic signals play an important role (6). Recently, the discovery that endothelial factors coming from aorta initiate pancreatic epithelium formation can explain that, until E9.5, vasculogenesis is the leading biological process that induct pancreas development; at E10.5, the dorsal pancreatic epithelium form a proper bud and after the separation of the aorta from the mesenchyme's layer, mesenchymal signals participate and this crosstalk becomes the guiding information to regulate pancreatic organogenesis through the successive embryonic stages. This information could drive the design of new drug. In order to characterize E10.5 stage, since it is the fundamental time point in which cells proliferate without neither differentiate nor initiate epithelial to mesenchymal transition, Laser Capture Microdissection (LCM): the identification of new molecules involved in pancreatic crosstalk can improve the simulation *in vitro* of cell differentiation biological processes (Figure 4).

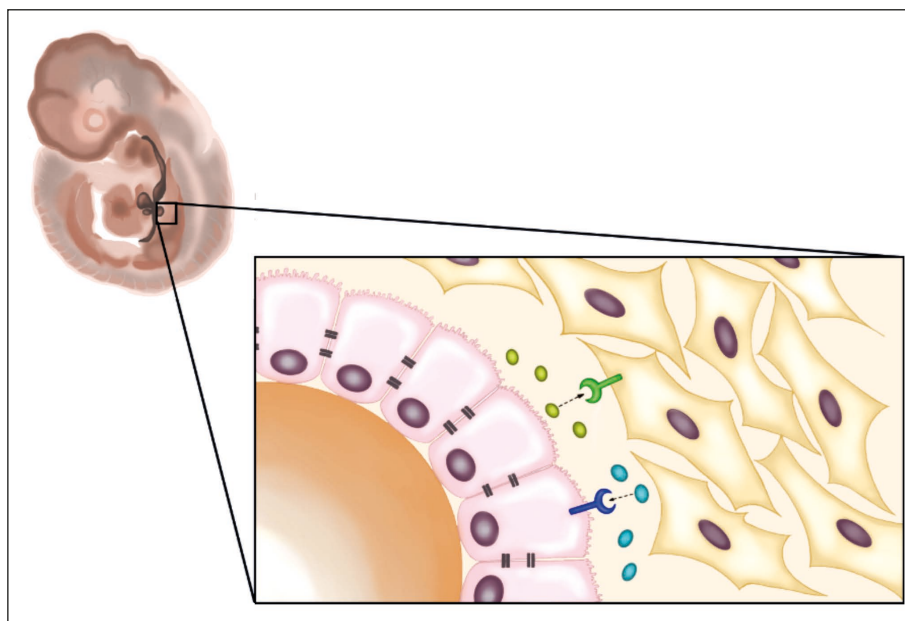


Fig. 4 - Schematic representation of mesenchyme and bud cellular cross-talk.

In conclusion, the evidences connecting pathologies and stem cell homeostasis are strong, and the underlying mechanistic bases linking stemness to tumor initiating events could foster novel clinical approach that could mimic native environment (7, 8).

Bibliografia essenziale

1. Sperka T, Wang J, Rudolph KL. DNA damage checkpoints in stem cells, ageing and cancer. *Nat Rev Mol Cell Biol.* 2012; 13: 579-590.
2. Bouwens L, Houbracken I, Mfopou JK. The use of stem cells for pancreatic regeneration in diabetes mellitus. *Nature reviews Endocrinology.* 2013; 9: 598-606.
3. De Angelis MT, Russo F, D'Angelo F, Federico A, Gemei M, Del Vecchio L, et al. Novel pancreas organogenesis markers refine the pancreatic differentiation roadmap of embryonic stem cells. *Stem cell reviews.* 2014; 10: 269-279.
4. Shih HP, Wang A, Sander M. Pancreas organogenesis: from lineage determination to morphogenesis. *Annual review of cell and developmental biology.* 2013; 29: 81-105.
5. Villasenor A, Cleaver O. Crosstalk between the developing pancreas and its blood vessels: an evolving dialog. *Seminars in cell & developmental biology.* 2012; 23: 685-92.
6. Gittes GK. Developmental biology of the pancreas: a comprehensive review. *Developmental biology.* 2009; 326: 4-35.
7. Harmon EB, Apelqvist AA, Smart NG, Gu X, Osborne DH, Kim SK. GDF11 modulates NGN3+ islet progenitor cell number and promotes beta-cell differentiation in pancreas development. *Development.* 2004; 131: 6163-6174.
8. Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. *Nature reviews Molecular cell biology.* 2014; 15: 786-801.

The role of RNA in epigenetic regulation of gene expression

Giuseppe Biamonti

Istituto di Genetica Molecolare del CNR, Pavia

For a long time, RNA molecules have been simply viewed as information-carrying intermediates in protein synthesis: the information stored in the gene sequence is transferred (transcribed) to a mRNA (messenger RNA) that is eventually translated into a protein.

Even in this simplistic system, however, two classes of non-coding RNA are required to get protein synthesis:

- 1) tRNAs, which represent a sort Rosetta stone to translate the nucleotide code into the aminoacid code;
- 2) ribosomes, i.e. large RNA-protein (ribonucleoprotein) assemblies that carry out protein synthesis. Forty years ago, the discovery of interrupted genes unveiled a more complex picture.

Now we know that, with a few exceptions, mammalian genes consist of stretches of protein-coding sequences called exons, separated by non-protein-coding regions called introns. Gene transcription produces long precursor RNAs (pre-mRNA) comprising both intronic and exonic sequences. Generation of the mature mRNA involves the precise removal of introns in a complex reaction called “splicing”. Notably, the large molecular machine involved in this reaction, i.e. the spliceosome, contains five small nuclear RNAs (snRNA).

Since then, the number of non-coding RNAs has drastically increased. Protein-coding sequences account for a mere 2% of the 3 billions of bases that compose the human genome. Even considering intronic sequences, genes represent about 55% of the genome.

The remaining part along with intronic sequences composes the “non-coding genome”, the function of which remains a major subject of study. Recently, deep sequencing analysis of the cell transcriptome revealed that most of the genome is actually transcribed. Although the function of most of the non-coding genome is still largely obscure, diverse classes of small and long non-coding (lnc) RNAs have been shown to play key roles in important biological processes such as regulation of gene expression, genome stability and protection against foreign genetic elements (1, 2).

Both small RNAs and lncRNAs can modify chromatin structure and silence transcription. Small RNAs act by guiding the recruitment of silencing complex-

es to target DNA sequences via a complementarity with nascent RNA molecule tethered to the DNA template by the transcriptional apparatus. The final step in this pathway is the recruitment of histone and DNA methyltransferases (3-5). Independently of small RNAs and RNA interference pathways, also chromatin-associated lncRNAs can recruit chromatin-modifying complexes (6, 7). In this short review I will discuss how non-coding RNAs contribute in establishing the epigenetic code and chromatin organization.

The “nascent transcript” model

Most of the data so far available about the involvement of small RNA in establishing heterochromatic domains derive from studies in the fission yeast in *S. pombe*. In that organism RNAi-mediated transcriptional gene silencing depends on the activity Argonaute 1 (Ago1), Dicer and RNA-dependent RNA polymerase. Inactivation of any of these functions results in loss of heterochromatic gene silencing at pericentromeric DNA repeats and by a reduced level of histone H3 lysine 9 (H3K9) methylation, a conserved heterochromatin marker (3). Heterochromatin formation relies on ability of the complex between a small RNA and Ago1 to recruit H3K9 methyltransferase Clr4 on specific chromosomal regions (8). According to the ‘nascent transcript’ model, chromatin-associated nascent RNAs would act as a scaffold for the assembly of large complexes composed of small-RNA and chromatin modifying factors (9, 10). In this system:

- 1) the nascent transcript is degraded by RNAi-dependent mechanisms;
- 2) RNAi-dependent H3K9 methylation leads to heterochromatin formation and transcriptional gene silencing (10, 11).

The basic features of the nascent transcript model, i.e. the recruitment of small RNA to chromatin-associated nascent RNAs with the ensuing assembly of large complexes appear to hold true also in other biological systems. Thus, similar strategy operates during DNA repair to mediate both gene silencing at DSBs and the assembly of DNA repair complexes in mammalian cells (12).

Long and short pericentromeric Satellite III RNA in human cells

The mechanism described above for the establishment of heterochromatin domains is specific of fungi and plants where a self-reinforcing positive feedback loop exists in which H3K9 methylation is dependent on the RNAi machinery and pericentromeric siRNA accumulation requires the H3K9 methyltransferase Clr4 (13). In animal cells, no conclusive evidence has been provided so far about the presence of small RNAs encoded by heterochromatic portions of the genome. We have found that different types of stress induce the transcription of constitutive heterochromatin. These RNAs derives from the massive transcription of long arrays of Satellite III repeats found at pericentromeric domains of specific human chromosomes, including chromosome 9 (14). Notably, SatIII RNA are eventually processed into a family of novel small RNAs that mark a specific period in the cell recovery from stress.

Piwi-interacting RNAs

Piwi-interacting RNAs (piRNAs) are small RNAs that interact with the largely germline-specific Piwi subfamily of Argonaute proteins. Although piRNAs are predominantly expressed in *Drosophila* germline cells, they have been detected in adult cells as well. Moreover, additional evidence indicates that piRNAs also appeared in human cancer cells (15). Their primary function is to silence the transposon elements and maintain the stability of the entire genome (16). The role of these molecules in guiding heterochromatin formation is suggested by the observation that artificial recruitment of Piwi to a reporter locus induces H3K9 methylation, HP1a accumulation and exclusion of RNA Pol II, suggesting that piRNAs may have a direct role in guiding chromatin changes (17). In mice, piRNAs silence transposons in the male germ line by targeting them for *de novo* DNA methylation during late embryonic and early neonatal development (18). This mechanism however, does not involve a self-reinforcing positive feedback loop observed in yeast. The molecular details underlying piRNA function in heterochromatin establishment are still to be clarified. In *Drosophila* piRNAs also contribute to epigenetic phenomena in which they are themselves responsible for epigenetic inheritance (19). Interestingly in *C. elegans* RNAi-induced H3K9 methylation is a transgenerational phenomenon, and the methylation pattern can be inherited for at least two generations through germline transmission of siRNAs (20).

Long non coding RNA trigger histone methylation without RNAi

LncRNAs, are defined as non-coding transcripts longer than 200 nucleotides. The number of these molecules is continuously expanding. Although their function is still matter of investigation (21), it is commonly accepted that many of these molecules act to target chromatin-modifying activities to particular genomic sites. An example is provided by lncRNAs involved in the process of dosage compensation in metazoans (22), such as the mammalian *XIST* RNA, which mediates inactivation of a randomly chosen X chromosome in females. It has been suggested that *XIST* functions by directing the Polycomb repressive complex 2 (PRC2), a H3K27 methyltransferase, to chromatin (23, 24). The idea that lncRNAs mediate heterochromatin assembly by interacting with PRC2 is still disputed mainly because PRC2 appears to bind RNAs in a very unspecific manner (25, 26). It has been suggested that additional factors may provide specificity to the complex. For example, the Jumonji family protein JARID2 was recently reported to act as an essential intermediate between *Xist* and PRC2 and JARID2 is required for the localization of PRC2 and methylation of H3K27 (24).

Enhancer lncRNAs

Interestingly a large number of mammalian enhancers are transcribed into lncRNAs, called enhancer RNAs (eRNAs), that seem to have major roles in pro-

moting the transcription of neighbouring genes. Little is known about the mode of action of these molecules. The HOXA distal transcript anti-sense RNA (*HOTTIP*) is an eRNA involved in the activation of HOXA homeobox genes. This eRNA induces chromosomal looping between the enhancer and gene promoters and recruits the H3K4 methyltransferase complex KMT2A to activate transcription of several genes in the HOXA cluster (27).

Thus, nascent RNA transcripts may act as scaffolds for the recruitment of co-activator complexes that mediate chromosome looping and transcriptional activation. Large RNAs (>200 nucleotides) may be particularly suited for such architectural tasks that bring enhancer and promoter regions, usually located great distances apart, into proximity.

Bibliografia

1. Moazed D. Small RNAs in transcriptional gene silencing and genome defence. *Nature*. 2009; 457: 413-420.
2. Cech TR, Steitz JA. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell*. 2014; 157: 77-94.
3. Volpe TA, Kidner C, Hall IM, et al. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science*. 2002; 297: 1833-1837.
4. Mochizuki K, Fine NA, Fujisawa T, Gorovsky MA. Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. *Cell*. 2002; 110: 689-699.
5. Gu SG, Pak J, Guang S, et al. Amplification of siRNA in *Caenorhabditis elegans* generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint. *Nat. Genet*. 2012; 44: 157-164.
6. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem*. 2012; 81: 145-166.
7. Bonasio R, Shiekhattar R. Regulation of transcription by long noncoding RNAs. *Annu. Rev. Genet*. 2014; 48: 433-455.
8. Verdel A, Jia S, Gerber S, et al. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science*. 2004; 303: 672-676.
9. Motamedi MR, Verdel A, Colmenares SU, et al. Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell*. 2004; 119: 789-802.
10. Bühler M, Verdel A, Moazed D. Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell*. 2006; 125: 873-886.
11. Bühler M, Haas W, Gygi SP, Moazed D. RNAi-dependent and -independent RNA turnover mechanisms contribute to heterochromatic gene silencing. *Cell*. 2007; 129: 707-721.
12. Michelini F, Pitchiaya S, Vitelli V, et al. Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. *Nat. Cell Biol*. 2017; 19: 1400-1411.

13. Noma K-I, Sugiyama T, Cam H, et al. RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nat. Genet.* 2004; 36: 1174-1180.
14. Biamonti G. Nuclear stress bodies: a heterochromatin affair? *Nat. Rev. Mol. Cell Biol.* 2004; 5: 493-498.
15. Mei Y, Clark D, Mao L. Novel dimensions of piRNAs in cancer. *Cancer Lett.* 2013; 336: 46-52.
16. Siomi MC, Sato K, Pezic D, Aravin AA. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* 2011; 12: 246-258.
17. Le Thomas A, Rogers AK, Webster A, et al. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* 2013; 27: 390-399.
18. Aravin AA, Sachidanandam R, Bourc'his D, et al. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell.* 2008; 31: 785-799.
19. de Vanssay A, Bougé A-L, Boivin A, et al. Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature.* 2012; 490: 112-115.
20. Ashe A, Sapetschnig A, Weick E-M, et al. piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell.* 2012; 150: 88-99.
21. Yang L, Froberg JE, Lee JT. Long noncoding RNAs: fresh perspectives into the RNA world. *Trends Biochem. Sci.* 2014; 39: 35-43.
22. Ferrari F, Alekseyenko AA, Park PJ, Kuroda MI. Transcriptional control of a whole chromosome: emerging models for dosage compensation. *Nat. Struct. Mol. Biol.* 2014; 21: 118-125.
23. Plath K, Fang J, Mlynarczyk-Evans SK, et al. Role of histone H3 lysine 27 methylation in X inactivation. *Science.* 2003; 300: 131-135.
24. da Rocha ST, Boeva V, Escamilla-Del-Arenal M, et al. Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. *Mol. Cell.* 2014; 53: 301-316.
25. Brockdorff N. Noncoding RNA and Polycomb recruitment. *RNA.* 2013; 19: 429-442.
26. Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat. Struct. Mol. Biol.* 2013; 20: 1250-1257.
27. Wang KC, Yang YW, Liu B, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature.* 2011; 472: 120-124.

Epigenetica e riprogrammazione cellulare

Ileana Zucchi

Istituto di Biotecnologie Biomediche, Consiglio Nazionale delle Ricerche, Segrate (Milano)

Cellular reprogramming

Cellular reprogramming is the epigenomic conversion of one specific cell type to another in both normal and disease development. In 2006, the field of stem cell biology was revolutionized when mouse embryonic fibroblasts (MEFs) were reprogrammed into induced pluripotent stem (iPS) cells (1) by the over-expression of four mouse transcription factors, Oct4, Sox2, Klf4, and c-Myc (mOSKM), using retroviral vectors, Figure 1.

During the generation of iPSCs, multiple changes have been observed with the MEFs, including changes in gene expression profiles (2, 3), epigenetic state (4, 5) cell morphology (6, 7) and cellular metabolism (8). Among these changes, the transition of mesenchymal to epithelial (MET) cell state has been recognized as a required step during the early phase of reprogramming (6, 7). Inhibiting early MET by inducing EMT with transformation growth factor (TGF)- β or Snail1 prevents iPSC generation.

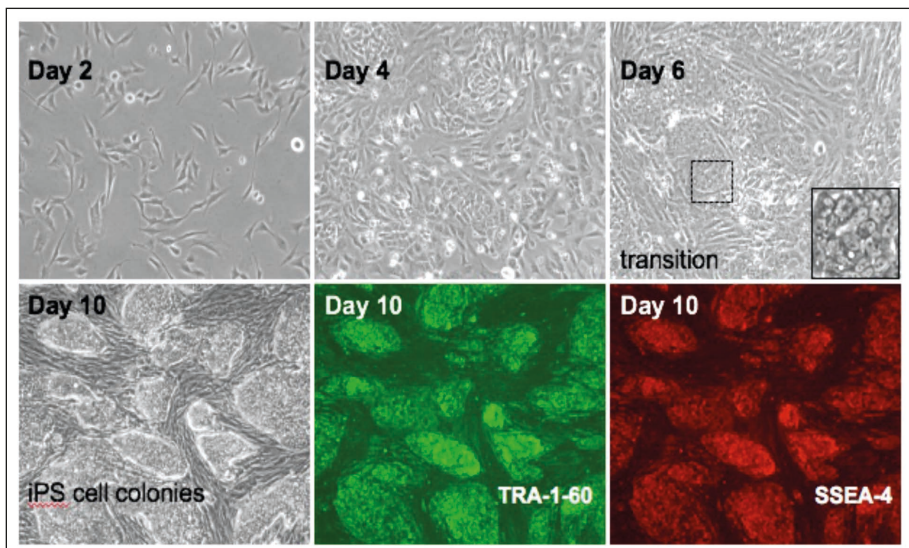


Fig. 1 - Reprogramming patient-derived fibroblast into iPSCs. (Kehler Zucchi 2017).

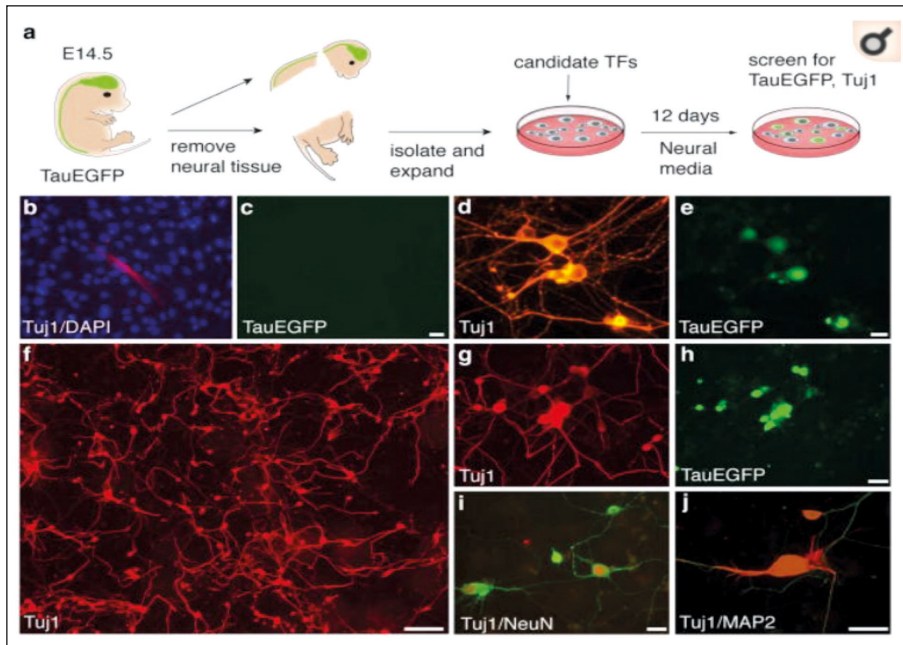


Fig. 2 - MEF-derived TuJ1-positive iN cells co-express the pan-neuronal markers TauEGFP (h), NeuN (red,i) and MAP2 (red,j). Vierbuchen et al., 2010.

More recently the direct reprogramming of fibroblasts to functional neurons was performed (9). The combination of only three factors, *Ascl1*, *Brn2* (also called *Pou3f2*) and *Myt1l*, was proven to be sufficient to rapidly and efficiently convert mouse embryonic and postnatal fibroblasts into functional neurons *in vitro*. These induced neuronal (iN) cells express multiple neuron-specific proteins, generate action potentials and form functional synapses, see Figure 2

The success of somatic cells to gain pluripotency after nuclear transfer or exogenous expression of four pluripotency-related transcriptional factors, and the direct reprogramming of fibroblasts into functional neurons make the iPSC and the trans-differentiation process a paradigm of cell fate conversion (2).

Cellular reprogramming during cell differentiation and Epithelial to Mesenchymal Transition (EMT) in normal development

Changing of one type of cells to another type occurs not only during somatic cell reprogramming but is crucial for adult stem cell differentiation and the generation of adult tissues and organs which require multiple rounds of sequential EMT and MET (10). As the first cell generated by a fertilized egg is a pluripotent epithelial cell that express Oct4, the generation of an embryo is dependent on EMT do generate mesenchyme cells for proper tissue and organ development (10-12). EMT is a cellular trans-differentiation process essential in normal embryo

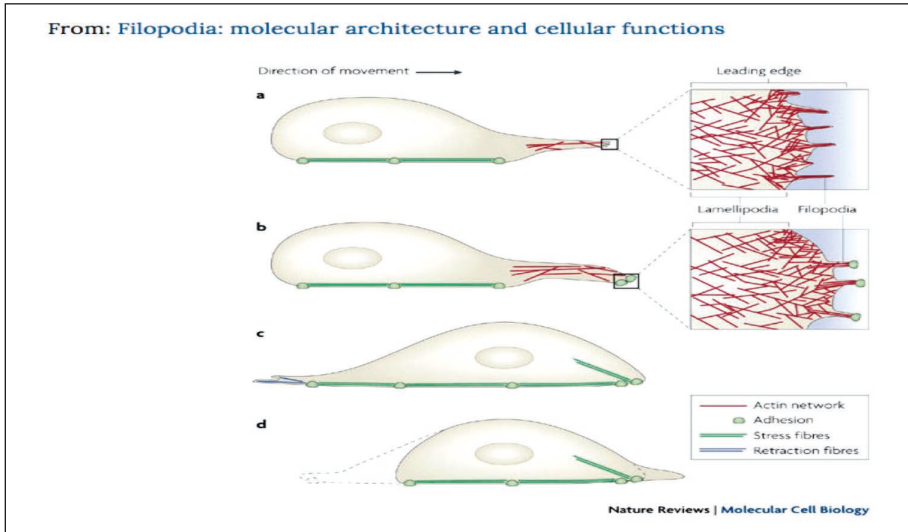


Fig. 3 - After undergoing EMT, cells acquire migratory and invasive properties that allow their invasion through the extracellular matrix (ECM).

development and wound healing (10-13), in which epithelial cells lose adhesion molecules and epithelial features to become mesenchyme-like cells and gain mobility capacity (14). The inter-conversion between epithelial cells and mesenchymal cells is a highly conserved and reversible cell process by which polarized, immotile epithelial cells extend filopodia from their basal surface and give rise to migrating mesenchymal cells, Figure 3.

Epigenetic regulators control cellular reprogramming

The activation of the EMT process depends on micro-environmental signals and on EMT-activating signals that interact with epigenetic regulators. These epigenetic regulators control the expression of proteins involved in several pathways, including: cell polarity, cell-cell adhesion, cell-cell contact, cytoskeleton and ECM degradation. In addition, these epigenetic regulators cause the repression of key epithelial genes and the acquisition of invasive properties and migratory capacities (15-20).

Epithelial and mesenchymal cells are two major types of cells in tissues and organs and the transitions between these two cell states as EMT and MET have been observed during multiple cell fate conversions including embryonic development (11). The development of heart has been considered as a good example for sequential EMT-MET. The formation of cardiac mesodermal cells during gastrulation and the organization of them into a two-layered epithelium later are considered as the first round of sequential EMT-MET. Another two rounds of sequential EMT-MET are observed during the folding around the primitive foregut and the formation of four heart compartments (12 and 21).

EMT in cancer and diseases

EMT is aberrantly activated in pathological conditions like tissue fibrosis and in other disease development such as sclerosis and tumor progression. Investigation into the mechanisms underlying cell fate conversions of MET-EMT may yield detailed molecular insights into cell fate decisions, not only for the switching between epithelial and mesenchymal cells, but also other cell types. Dissecting molecular pathways and components associated with normal and disease EMT is also critical to developing therapeutic strategies against EMT-associated diseases including cancer, aberrant cell dissemination and metastasis.

The EMT has been proposed as a preliminary event underlying the metastatic process, allowing tumor cells to acquire skills in migration and colonization of new tissues and organs (17). Cell undergoing EMT lose apico-basal polarity, intercellular and cell-extra-cellular adhesions and expression of epithelial makers, gain fibroblast-like morphology and expression of fibroblast markers, reorganize the cytoskeleton and acquire cell motility and invasive behaviour (17) allowing the tumor cells to invade, migrate into the circulatory system and circulate into blood vessels (Figure 4).

MicroRNA in the regulation of EMT associated to cancer invasion

MicroRNAs are small (20-30 nucleotides) noncoding RNAs that typically inhibit the translation and stability of messenger RNAs (mRNAs) controlling gene activity involved in diverse cellular processes such as inflammation, cell-cycle regulation, stress response, differentiation, apoptosis and cell migration (22, 23). While involvement of microRNAs (miRs, miRNAs) in the regulation of EMT processes and their dysregulation in cancer invasion and progression is well estab-

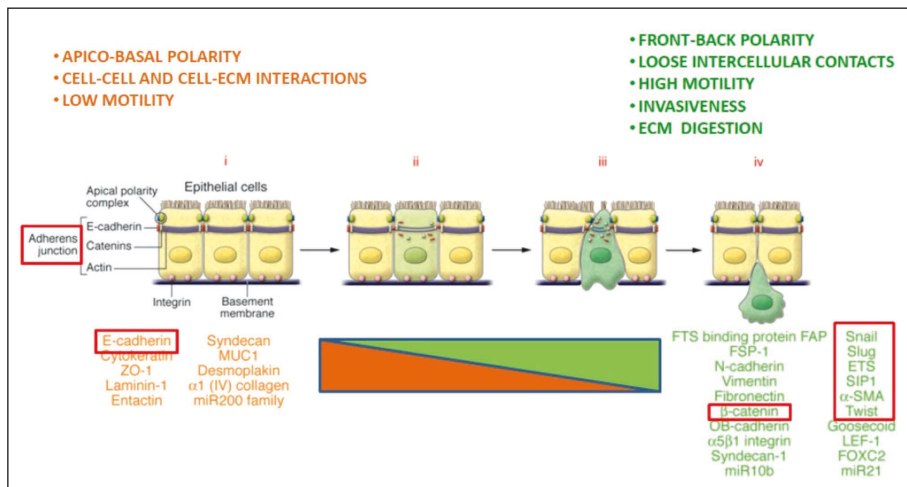


Fig. 4 - Epithelial-to-mesenchymal transition (EMT) is a process in which epithelial cells lose their epithelial identity to gain a motile mesenchymal phenotype.

lished (22, 23), the target genes of EMT associated miRs in the development and progression of cancer are still largely unknown (22, 23).

Our group has developed a model of mammary carcinogenesis based on the rat cancer stem-like cells (CSCs) LA7, that have the capacity to generate all cell types of the mammary parenchyma and 3D tubulo-alveolar architecture in vitro and in vivo (24). LA7, when engrafted as a single cell into NOD-SCID mice, generate a primary tumor that contains self-renewing CSCs, luminal and myoepithelial differentiated cells and cells at various stages of epithelial differentiation. Additionally, during tumor development, the CSC or its progeny generates elongated-mesenchyme-like cells expressing EMT markers, (24, 25). As single LA7CSC-like cells can generate both mammary luminal cells and mesenchyme like cells, LA7CSCs are an ideal model to study EMT and MET as the single cell level (Figure 5).

The LA7CSC model recapitulates the entire process of tumor development, including invasion into the extracellular matrix (ECM), generation of EMT cells with migration capacity to blood vessels and invasion into the blood stream resulting in

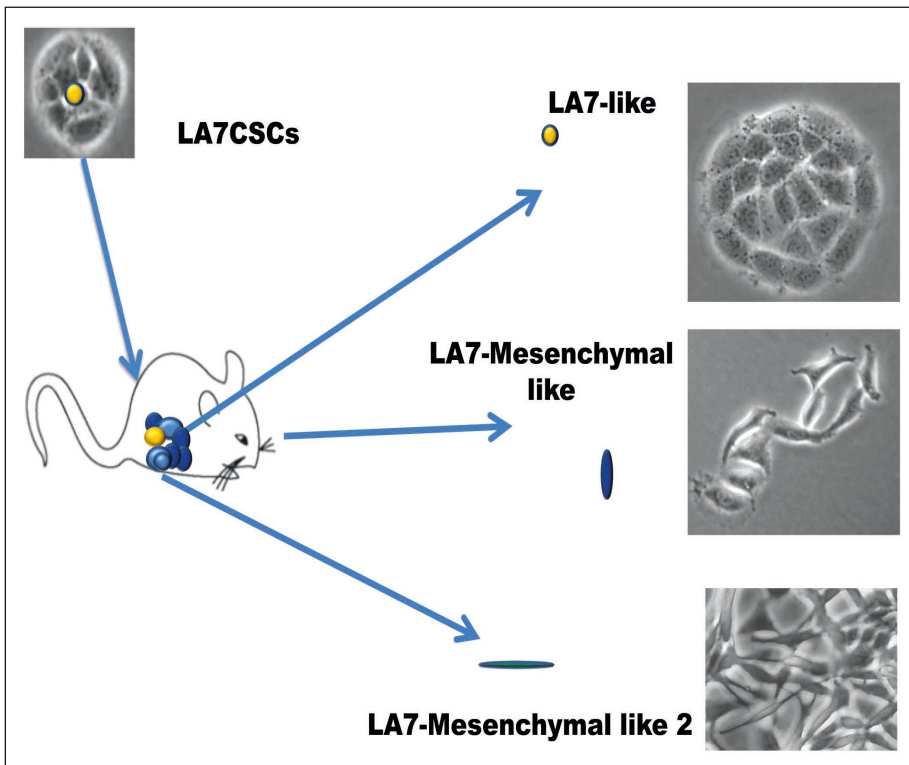


Fig. 5 - A single LA7 cell engrafted in NOD-SCID mice, generate a primary tumor that contains self-renewing CSCs (LA7-like), similar in every way to the LA7 cell transplanted, and a progeny of differentiated cells and elongated-mesenchymal-like cells. In 100 days, the single LA7 generates secondary tumors to distant sites through a metastatic process requiring EMT and MET.

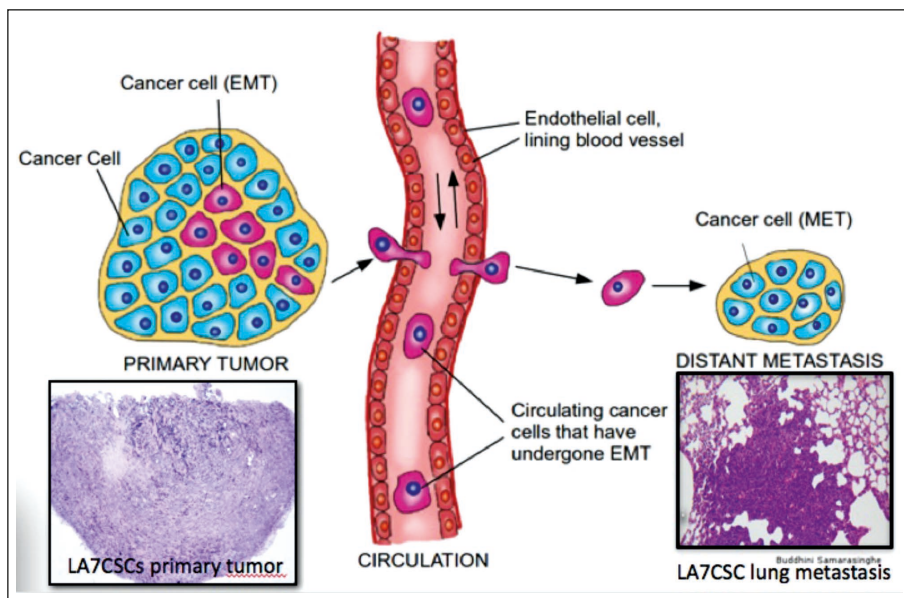


Fig. 6 - LA7CSCs are involved in all the stages of tumor development, progression and metastasis.

the circulation of cancer cells with SC properties and metastasis formation in for instance the liver and kidney (25), Figure 6. The secondary tumors generated have a morphology identical to the primary, consisting of tubule-alveoli 3D structures suggestive of a conversion of cells generated by EMT back to MET cells.

The LA7CSC model recapitulates the entire process of tumor development, including invasion, migration to the blood stream, circulation of cancer cells in blood system and metastasis formation.

Since the elongated mesenchyme like cells are rat cells and can be differentiated from the resident mesenchyme cells of the mouse used for the xeno-transplantation study, the rat mesenchyme like cells are therefore generated from the single CSC transplanted or progeny cells generated from the single rat CSC. We therefore consider that the elongated mesenchyme like cells were generated by an epigenomic event by which a single LA7CSC or its progeny underwent a change in cell identity. To dissect the molecular pathways associated with EMT and MET, and their associated metastasis-promoting-function we performed microarray- and micro-RNA-array screening and a proteomic approach based on SILAC. (Zucchi Unpublished).

Our results show that network of microRNA drives epithelial and mesenchyme like cell reprogramming and tumour progression. We determined a set of 16 miR regulates the induction of EMT or MET by repressing transcription of epithelial, or mesenchyme genes respectively. These miRs have a key role in cell fate determination and maintenance by regulating mRNAs associated with epigenetic changes in the dynamics of chromatin, alternative splicing, post-transla-

tional modifications of histones and remodelling and maintenance of the nuclear architecture.

Collectively our results show that the regulation of EMT and MET in LA7CSs and their progeny are associated with epigenetic reprogramming regulated by microRNAs.

Bibliografia

1. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006; 126: 663-676.
2. Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007; 131: 861-872.
3. Brambrink T, Foreman R, Welstead GG, et al. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell*. 2008; 2: 151-159.
4. Koche RP, Smith ZD, Adli M, et al. Bernstein BE, Meissner A. Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell*. 2011; 8: 96-105.
5. Watanabe A, Yamada Y, Yamanaka S. Epigenetic regulation in pluripotent stem cells: a key to breaking the epigenetic barrier. *Philos Trans Royal Soc London Ser B Biol Sci*. 2013; 368: 20120292.
6. Li R, Liang J, Ni S, et al. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell*. 2010; 7: 51-63.
7. Samavarchi-Tehrani P, Golipour A, David L, et al. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell*. 2010; 7: 64-77.
8. Folmes CD, Nelson TJ, Martinez-Fernandez A, et al. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metabol*. 2011; 14: 264-271.
9. Vierbuchen T, Ostermeier A, Pang ZP, et al. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*. 2010; 463: 1035-1041.
10. Boiani M, Gentile L, Gambles V, Cavaleri F, Redi CA, Schöler HR (2005) Variable "reprogramming" of the pluripotent stem cell marker Oct4 in mouse clones: distinct developmental potentials in different culture environments. *Stem Cells*. 2005; 23: 93-108.
11. Nakaya Y, Sheng G. Epithelial to mesenchymal transition during gastrulation: an embryological view. *Devel Growth Differ*. 2008; 50: 755-766.
12. Thiery JP, Acloque H, Huang RY, et al. Epithelial-mesenchymal transitions in development and disease. *Cell*. 2009; 139: 871-890.
13. Kalluri R, Neilson EG. Epithelial-mesenchymal transition and its implications for fibrosis. *J. Clin. Invest*. 2003; 112: 1776-1784.
14. Nieto MA. The ins and outs of the epithelial to mesenchymal transition in health and disease. *Annu. Rev. Cell Dev. Biol*. 2011; 27: 347-376.

15. Puisieux A, Brabletz T, Caramel J. Oncogenic roles of EMT-inducing transcription factors. *Nat. Cell Biol.* 2014; 16: 488-494.
16. Grigore AD, Jolly MK, Jia D, et al. Tumor Budding: The Name is EMT. *Partial EMT. J. Clin. Med.* 2016; 5: E51
17. Lee JM, Dedhar S, Kalluri, R, et al. The epithelial-mesenchymal transition: New insights in signaling, development, and disease. *J. Cell Biol.* 2006; 172: 973-981.
18. Valastyan, S. Weinberg, R.A. Tumor metastasis: Molecular insights and evolving paradigms. *Cell.* 2011; 147: 275-292.
19. Nieto MA, Huang R, Jackson. EMT: *Cell* 2016; 166, 21-45.
20. Brabletz T, Kalluri R, Nieto AM, et al. EMT in Cancer. *Nature Reviews Cancer.* 2018; 18: 128-134.
21. Nakajima Y, Yamagishi T, Hokari S, et al. Mechanisms involved in valvuloseptal endocardial cushion formation in early cardiogenesis: roles of transforming growth factor (TGF)-beta and bone morphogenetic protein (BMP) *Anat Rec.* 2000; 258: 119-127.
22. Zaravinos The Regulatory Role of MicroRNAs in EMT and Cancer. *J Oncol.* 2015; 865816.
23. Drago-Garcia D, Espinal-Enriquez J, Hernandez-Llemus E. Network analysis of EMT and MET micro-RNA regulation in breast cancer *Scientific Reports.* 2017; doi:10.1038/s41598-017-13903-1
24. Zucchi I, Sanzone S, Astigiano S, et al. The properties of a mammary gland cancer stem cell. *Proc Natl Acad Sci USA.* 2007; 104: 10476-10481.
25. Zucchi I, Astigiano S, Bertalot, et al. Distinct populations of tumor-initiating cells derived from a tumor generated by rat mammary cancer stem cells. *Proc Natl Acad Sci USA.* 2008; 105: 16940-16945.
26. Chen T, Dent SYR. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat Rev Genet.* 2014; 15: 93-106.
27. Skrypek N, Goossens S, De Smedt, E, Vandamme N, Berx G. Epithelial-to-mesenchymal transition: epigenetic reprogramming driving cellular plasticity. *Trends Genet.* 2017; 33, 943-959

Metodologie di indagine molecolare e strutturale applicate all'identificazione delle modificazioni post-traduzionali

Federico Forneris

The Armenise-Harvard Laboratory of Structural Biology, Department of Biology and Biotechnology, University of Pavia

Questo testo non ha la pretesa di offrire una accurata e comprensiva panoramica sulle strategie da adottare per caratterizzare le modificazioni post-traduzionali. L'obiettivo che si vuole raggiungere è quello di stimolare, in lettori interessati, la volontà di approfondire, prima di tutto attraverso le fonti citate e più in generale nella letteratura scientifica contemporanea, le metodologie disponibili per affrontare lo studio di una tematica biologica che al momento è largamente approfondita ma nonostante tutto ancora molto complessa da analizzare sotto il profilo molecolare. Buona lettura.

Introduzione

Il significato biologico delle modificazioni post-traduzionali

Molti processi biologici sono regolati attraverso modifiche capaci di attivare o di inattivare le proteine e gli enzimi in essi coinvolti. Questi processi di regolazione spesso avvengono a livello trascrizionale, andando ad alterare selettivamente l'espressione dei geni coinvolti e quindi riducendo (o aumentando) anche di vari ordini di grandezza, la quantità di proteine codificate da questi geni (1, 2). Una seconda strategia, largamente presente nei sistemi biologici e che consente uno spettro molto ampio di funzionalità, agisce direttamente sulle proteine, alterandone o regolandone funzioni specifiche e architetture molecolari. Queste modificazioni, dette per la loro natura "post-traduzionali", rappresentano quindi una risorsa importante per controllare l'omeostasi all'interno e all'esterno delle cellule negli organismi viventi. Molti eventi di regolazione epigenetica, a livello molecolare, avvengono a seguito dell'introduzione, della modificazione o della rimozione di specifiche modificazioni post-traduzionali (PTM). Si pensi ad esempio alle numerose modificazioni cui possono essere soggette le proteine istoniche (1, 2): la varietà e i numerosi significati biologici associati a questi eventi hanno portato addirittura all'elaborazione di un vero e proprio "codice istonico", all'interno del quale specifiche PTM vengono associate a precisi eventi di attivazione o di repressione genica nei contesti più diversificati, dal differenziamento tissutale

all'ematopoiesi, alla neurogenesi, nonché processi legati allo sviluppo e alla progressione di malattie tumorali (1, 3, 4).

A livello molecolare, le PTM possono alterare, anche in modo piuttosto marcato, il network di interazioni specifiche di una proteina. In particolare, l'introduzione o la rimozione di modificazioni su catene laterali di specifici aminoacidi può modificare la propensione ad instaurare interazioni di carattere elettrostatico e/o idrofobico, indurre stabilizzazione in specifici ripiegamenti proteici (talvolta generando nuove piattaforme per contatti intermolecolari), oppure facilitare la creazione di segmenti proteici molto flessibili in regioni altrimenti rigide e ben definite. Queste alterazioni possono coinvolgere un'area molto localizzata intorno all'aminoacido soggetto alla modificazione, oppure propagare i loro effetti a tutta la struttura proteica a causa di alterazioni sistematiche dell'insieme di legami a idrogeno e di interazioni idrofobiche anche a lunga distanza dalle modificazioni stesse. In casi come questi, la modificazione può diventare così pronunciata da indurre transizioni conformazionali anche molto estese, portando a spostamenti di interi domini proteici e causando vere e proprie "trasformazioni" molecolari (si veda (5) per un esempio legato ad eventi di proteolisi nel sistema immunitario innato).

È chiaro che modificazioni di questa entità a livello molecolare si traducono, in contesti cellulari e tissutali in eventi che possono essere anche molto significativi. Si pensi al caso delle cascate di attivazione delle chinasi: l'aggiunta di un gruppo fosfato su un enzima lo altera in modo così significativo da renderlo capace di legarsi con alta affinità ad un'altra proteina e modificarla; in assenza di fosforilazione, l'affinità è così bassa da prevenire qualsiasi tipo di riconoscimento (1, 2). L'attivazione consecutiva di più chinasi porta a processi di signaling cellulare che possono determinare il destino della cellula stessa, attivando o inibendo processi di replicazione, differenziamento, o apoptosi e morte cellulare. All'interno del nucleo delle cellule, PTM epigenetiche sulle code proteiche degli istoni alterano il grado di compattazione della cromatina, facilitando o impedendo l'espressione genica a causa di una ridotta accessibilità delle sequenze di DNA da trascrivere (4).

Oltre a ruoli funzionali essenziali, le PTM hanno anche importanti ruoli di carattere strutturale: l'introduzione di modifiche su proteine rappresenta infatti un sistema molto efficiente per rendere più stabili e rigide queste macromolecole. La cromatina è nuovamente un esempio calzante: l'introduzione di molte PTM epigenetiche repressive causa una forte compattazione che da un lato riduce l'accessibilità per i processi trascrizionali, dall'altro conferisce quella stabilità e rigidità che caratterizza regioni cromosomali ben definite ed essenziali quali telomeri e centromeri (6-8). Nello spazio extracellulare, fondamentali sono le PTM sulle molecole del collagene, modifiche che costituiscono una percentuale significativa di queste molecole essenziali per la formazione di tessuti e organi: senza PTM i collagene non possono acquisire quelle particolari caratteristiche di elasticità e resistenza che caratterizzano il tessuto connettivo e consentono la creazione di quel complesso network di macromolecole che è la matrice extracellulare (9).

Un affascinante aspetto legato alle PTM è la forte correlazione tra i siti di possibile modificazione e la presenza di regioni disordinate all'interno del folding proteico: regioni molto flessibili e prive di strutture secondarie ben definite presentano infatti un'accentuata propensione ad adattarsi per interagire con molteplici partners molecolari, nonché una forte propensione all'essere siti per PTM (10). Non stupisce dunque notare che le proteine che presentano maggiori caratteristiche di flessibilità siano anche quelle che presentano più frequentemente PTMs. È il caso ad esempio della studiatissima p53, una proteina dai molteplici ruoli associati allo sviluppo e alla proliferazione di molti tumori. p53 alterna nella sua sequenza aminoacidica regioni strutturalmente organizzate e ben definite ad elementi molto flessibili e disordinati, molti dei quali sono siti di modificazioni post-traduzionali che presentano ruoli critici nei meccanismi in cui p53 è coinvolta. Queste modificazioni determinano il reclutamento di partner specifici e possono alterare drasticamente il significato biologico (e quindi il destino) delle cellule in p53 modificata è presente (11). Analogamente, le già menzionate code N-terminali degli istoni costituiscono un altro esempio paradigmatico di come flessibilità e PTM siano fortemente correlate (3, 4).

Studiare le PTM a livello molecolare

La natura spesso dinamica dei processi che portano all'aggiunta o alla rimozione delle PTM, unita alla flessibilità conformazionale appena menzionata, sono chiari indici di quanto sia complesso riuscire ad identificare e caratterizzare queste importanti modificazioni. Le numerose implicazioni biologiche tuttavia giustificano sforzi anche considerevoli per poter raggiungere una comprensione molecolare accurata e, laddove possibile, la visualizzazione della presenza delle PTM associate a specifici eventi della vita della cellula. Una difficoltà aggiuntiva da non trascurare quando si vogliono studiare PTM è associata alla natura stessa di queste modificazioni: il ricorso sempre più frequente a sistemi di produzione eterologa per caratterizzare sistemi proteici può non essere compatibile con studi di questo tipo, in quanto non è detto che meccanismi di produzione ricombinante siano in grado di riprodurre fedelmente gli apparati di modificazione post-traduzionale endogeni. Questo è ovviamente valido quando si fa ricorso a sistemi ricombinanti molto più semplici rispetto a quelli oggetto di studio (es. produzione di proteine umane in cellule batteriche), ma lo è altrettanto anche quando si analizzano tipologie cellulari differenti nella stessa specie (es. produzione di proteine umane in cellule umane differenti da quelle in cui le proteine oggetto di studio svolgono le loro funzioni).

La strategia convenzionalmente adottata per identificare le PTM dal punto di vista molecolare è il western blotting (WB). Questa tecnica si basa sul riconoscimento specifico, da parte di anticorpi, della PTM presenti sulle proteine (o del sito in cui dovrebbe essere presente, però privo della modificazione). Basandosi sui meccanismi di interazione molecolare tra antigene (il sito della modificazione in questo caso) e anticorpo, ed essendo una tecnica accessibile e di facile implementazione in pressoché qualsiasi laboratorio di biologia molecolare, il WB può

a pieno titolo essere considerato una delle principali risorse per lo studio delle PTM. Un ulteriore vantaggio di questo metodo è che, con le dovute accortezze, può essere utile anche per analisi di tipo quantitativo e non solo qualitativo (12). Tuttavia, è importante tenere presente che questa tecnica dipende fortemente dalla qualità e dalla affidabilità dei reagenti (anticorpi) utilizzati per le indagini e da come questi sono stati generati. I processi di immunizzazione richiedono antigeni ad elevato grado di purezza ed omogeneità, che nel caso di campioni con PTM potrebbero non essere così facili da reperire: si crea dunque un circolo vizioso in cui il reagente per l'identificazione della PTM dipende dalla PTM stessa, una configurazione sperimentale chiaramente non ottimale. Per questa ragione, è fondamentale poter fare affidamento almeno su una seconda metodologia di indagine, in modo da ottenere un controllo incrociato sulla qualità e sul significato delle informazioni raccolte. Tipicamente, questo obiettivo viene raggiunto mediante indagini a livello cellulare (caratterizzando specifici eventi dovuti alla presenza delle modificazioni quali ad esempio variazioni nei fenotipi cellulari o attivazione di cascate di differenziamento o apoptosi), oppure mediante indagini su processi molecolari che si innescano a seguito della PTM, come ad esempio eventi di fosforilazione associati a cascate di kinasi o attivazione/repressione di processi di trascrizione genica. In entrambi questi casi, risulta evidente che, come già descritto per il caso degli anticorpi, è necessaria una precisa conoscenza pregressa di ciò che si verifica all'interno della cellula in presenza di una determinata PTM.

Ma allora come si possono studiare "nuove" modificazioni? È necessario fare ricorso a metodologie di indagine che non dipendano dalle modificazioni stesse, capaci quindi di identificare l'eventuale presenza di queste modificazioni anche in assenza di qualsiasi conoscenza pregressa in merito. Considerato il funzionamento delle metodiche convenzionalmente utilizzate per questa tipologia di studi, è chiaro che quello che può sembrare un discorso assolutamente triviale non lo è per niente.

Un metodo che soprattutto negli ultimi anni ha davvero rivoluzionato l'identificazione delle PTM è la spettrometria di massa (MS). L'utilizzo di MS alle indagini molecolari non è nuova: procedure quali la determinazione di sequenze aminoacidiche da digeriti triptici sono da più di 20 anni alla base di studi di proteomica (13). Questi metodi si basano sulla determinazione di picchi associati ai rapporti massa/carica (m/z) dei peptidi ottenuti dalla digestione enzimatica dei campioni proteici: i valori di m/z ottenuti corrispondono in modo specifico alla sequenza aminoacidica presente all'interno dei peptidi stessi. Tramite una indagine MS accurata condotta in questo modo è possibile quindi identificare intere sequenze aminoacidiche, anche in matrici proteiche complesse. La presenza di PTM complica in modo significativo l'indagine MS, poiché va ad alterare i valori di m/z degli aminoacidi soggetti a modificazione. Negli ultimi anni sono stati introdotti sul mercato strumenti MS ad alta risoluzione, capaci di risolvere spettri di massa anche molto complessi su intervalli di m/z molto ampi. Unitamente allo sviluppo di softwares sempre più avanzati per l'assegnazione automatica dei picchi sperimentali a precise sequenze aminoacidiche (tenendo conto anche di possibili PTM), questi strumenti costituiscono un notevole passo avanti per le

indagini sulle PTM, poiché permettono di determinare sperimentalmente, anche all'interno di campioni ignoti, la possibile presenza di PTM a carico di specifici residui aminoacidici. L'alta risoluzione è fondamentale: da un lato, permette di identificare anche piccole variazioni di massa (es. una idrossilazione +16 Da) su proteine intere di grandi dimensioni (decine di kDa); dall'altro, permette di risolvere ed assegnare con precisione picchi di m/z in matrici anche molto complesse, che generano spettri con picchi spesso convoluti. Inoltre, un importante valore aggiunto legato ai più avanzati strumenti MS ad alta risoluzione è legato alla possibilità di valutare l'abbondanza relativa di picchi anche molto simili tra loro, permettendo ad esempio di valutare la percentuale di PTM presente su di un determinato aminoacido in un peptide rispetto alla quantità totale di peptide presente (13). Questo permette, analizzando in sequenza più campioni trattati in modo differente, non solo di identificare nuove PTM, ma anche di stabilire come la loro comparsa (o scomparsa) vari in funzione di tempi e modalità specifiche di trattamento del campione. Insomma, un importante valore aggiunto nelle moderne indagini di proteomica.

Una volta identificate e quantificate le PTM, l'indagine molecolare può spesso richiedere la valutazione dei meccanismi associati alla loro introduzione/rimozione e/o la comprensione delle strutture tridimensionali delle macromolecole in cui queste sono localizzate. Per quanto riguarda gli aspetti meccanicistici, questi sono ovviamente soggetti ad estrema variabilità a seconda del sistema molecolare oggetto di studio. Considerato l'ampio spettro di modificazioni e di sistemi biologici coinvolti nella loro introduzione e modificazione, un ulteriore approfondimento in questa sede sarebbe eccessivo e pertanto si invita alla consultazione di più specifici testi di biochimica e biologia molecolare (1, 2). Per quanto riguarda invece l'analisi strutturale, è possibile affrontare questo tipo di tematiche mediante un discorso più generale che può essere applicato a molteplici sistemi molecolari, anche molto diversi tra loro.

Al giorno d'oggi gli studi strutturali legati alle PTM si fondano sostanzialmente su due metodologie di indagine:

- 1) studi su campioni in soluzione, nei quali è possibile valutare non soltanto gli aspetti specifici legati alle interazioni tra le PTM e l'intorno proteico, ma anche le implicazioni legate alla flessibilità conformazionale spesso associata alle regioni soggette a PTM;
- 2) studi su campioni intrappolati in matrici solide (cristalli proteici, ghiaccio vitrificato), che consentono spesso di raggiungere risoluzioni atomiche ma che rappresentano una soluzione di compromesso per quanto riguarda la comprensione degli aspetti di dinamicità e flessibilità conformazionale.

L'indagine strutturale delle PTM in soluzione

La tecnica principale per gli studi strutturali in soluzione è la Risonanza Magnetica Nucleare o NMR (14). Attraverso l'NMR, è possibile ottenere una mappatura delle distanze tra i residui aminoacidici che costituiscono il polipeptide, e da questa mappatura ricostruire quindi la macromolecola attraverso procedure

di triangolazione. Quando si studiano PTM l'NMR può essere di grande aiuto nell'indagine strutturale, permettendo di comparare il campione in soluzione con e senza modificazioni. Questo tipo di indagine consente di mettere in evidenza le variazioni a carico dei residui che interagiscono direttamente con la PTM, generando quindi una mappa accurata di ciò che accade a livello strutturale in presenza o in assenza della modificazione. Questa analisi assume particolare importanza quando si valutano PTM localizzate in regioni flessibili, poiché consente di mettere in evidenza possibili contatti intramolecolari causati dalla presenza/assenza della modificazione. Qualora si vogliano invece mettere in evidenza contatti di tipo intermolecolare formati a seguito di PTM, è necessario fare ricorso a metodologie avanzate di *labeling* isotopico, in cui si va ad "accendere" selettivamente il segnale NMR di un solo componente proteico (tipicamente quello con le PTM) e se ne valutano le variazioni a seguito dell'aggiunta di possibili interattori. Variazioni nei segnali in prossimità della PTM indicheranno possibili contatti intermolecolari con gli interattori che sono stati aggiunti nell'esperimento.

Un approccio differente per analizzare gli effetti delle PTM in soluzione sfrutta sistemi di *labeling* isotopico simili a quelli appena descritti per l'NMR e li combina con l'indagine attraverso la spettrometria di massa. È il caso del *hydrogen-deuterium exchange mass spectrometry* o HDX, un approccio molto efficace per studiare contatti intermolecolari. Tramite questo metodo è possibile individuare quali porzioni specifiche di una superficie di una macromolecola sono coinvolte in una interazione e se queste comprendono regioni affette da PTM. Questo metodo presenta una minor complessità rispetto al *labeling* per NMR ed è per questo molto utilizzato come strumento integrativo nelle analisi strutturali: ad esempio, lo sfruttamento di strutture cristallografiche (che mostrano conformazioni "rigide") di due proteine soggette a modificazioni e la loro analisi HDX in soluzione può fornire una panoramica completa delle interazioni che si vengono a formare a seguito di PTM, facilitando quindi la comprensione degli eventi molecolari ad esse associati (15).

L'indagine strutturale delle PTM in matrici solide

L'approccio più noto e più convenzionale per caratterizzare le strutture molecolari è la diffrazione di raggi X su cristalli di proteine, chiamata anche più semplicemente cristallografia a raggi X. La sua applicazione allo studio delle PTM richiede ovviamente la possibilità di intrappolare la macromolecola soggetta a PTM all'interno di un cristallo, per poterne poi determinare la struttura tridimensionale attraverso procedure matematiche di analisi delle immagini di diffrazione che questo genera una volta esposto a raggi X (14). Requisito essenziale per uno studio accurato su PTM è che le modificazioni si collochino in regioni della macromolecola non soggette a variabilità conformazionale, così da poterne ottenere una interpretazione strutturale statica ma molto accurata per quanto riguarda le interazioni che coinvolgono le PTM. Un vantaggio importante della cristallografia rispetto alle tecniche in soluzione è che questa metodica è applicabile sia a macromolecole isolate che a complessi, purché questi possano essere cristallizza-

ti. Lo svantaggio più significativo riguarda invece la caratterizzazione di regioni flessibili, impossibili da intrappolare all'interno di un rigido ambiente cristallino. Come già discusso, non va esclusa tuttavia la possibilità che una regione anche molto flessibile possa, a seguito di PTM, diventare molto rigida a seguito di interazioni specifiche: in casi come questo lo studio cristallografico sarà sicuramente possibile per le conformazioni rigide, e potrà fornire informazioni cruciali in merito alle interazioni a ridosso delle PTM.

Un trend metodologico che negli ultimi anni sta veramente rivoluzionando la biologia strutturale è quello legato alla (crio)-microscopia elettronica (16, 17). Questi metodi consentono di ottenere strutture tridimensionali di macromolecole biologiche senza dover ricorrere alla cristallizzazione, che spesso rappresenta il passaggio più complicato di una caratterizzazione strutturale a risoluzione molecolare. L'applicazione della microscopia elettronica allo studio delle PTM tuttavia non è immediato, in quanto questa tecnica è al momento in grado di fornire risultati eccezionali solo per quanto riguarda macromolecole di grandi dimensioni (con masse al di sopra dei 200 kDa), e presenta limiti al momento per quanto riguarda il raggiungimento di risoluzioni elevate, indispensabili per visualizzare la presenza di specifiche alterazioni sulle catene laterali di aminoacidi. Considerando però gli effetti "macroscopici" che le PTM possono generare (modificazioni nell'organizzazione di interi domini, alterazioni nel riconoscimento molecolare di interattori), è chiaro che l'applicazione di questi moderni metodi di indagine può rappresentare, anche per lo studio delle PTM, un passo avanti notevole. Se si considera inoltre che la crio-microscopia elettronica sembra essere un metodo eccezionalmente adatto per lo studio di complessi macromolecolari contenenti nucleosomi e cromatina (18), è intuitivo immaginare come questi metodi potranno, nel futuro prossimo, fornire informazioni ragguardevoli e significative sui meccanismi biologici associati a molte PTM su istoni, nucleosomi e proteine ad essi associati.

Conclusioni

L'indagine molecolare sulle PTM rappresenta al giorno d'oggi una importante risorsa per comprenderne il loro funzionamento e le implicazioni a livello di organi e tessuti. L'implementazione di metodiche per l'identificazione qualitativa e quantitativa, combinate con i più moderni approcci di visualizzazione strutturale, definiscono lo stato dell'arte nella comprensione delle implicazioni associate a queste piccole variazioni dal grande significato biologico. È interessante notare come metodiche che ad un primo impatto potrebbero non apparire utili allo scopo (es., lo studio di regioni intrinsecamente disordinate mediante cristallografia, oppure l'identificazione di una minuscola PTM in una complessa mappa strutturale ottenuta mediante crio-microscopia elettronica) possono invece rivelarsi indispensabili proprio a causa delle alterazioni che precise PTM possono indurre nelle macromolecole (es., stabilizzando una regione disordinata in un complesso macromolecolare o introducendo una variazione conformazionale estesa su domini proteici). Un aspetto davvero intrigante di questo tipo di studi è associato alla rapida evoluzione che queste tecniche stanno subendo negli ultimi anni: meno di un

decennio fa l'individuazione e la caratterizzazione delle PTM rappresentava una impresa quasi titanica, mentre al giorno d'oggi gli strumenti per proteomica spesso offrono l'individuazione delle PTM come un valore aggiunto quasi "gratuito" all'interno di caratterizzazioni estese. La possibilità poi di sfruttare più metodiche di indagine, qualitative e anche quantitative, per localizzare e comprendere le PTM consente un vero e proprio approccio ortogonale di validazione molecolare delle PTM osservate, che ne facilita da una parte la corretta individuazione e nel contempo consente di verificare informazioni pubblicate negli anni passati.

Bibliografia

1. Alberts B. *Molecular biology of the cell*. 4th ed. New York: Garland Science; 2002.
2. Berg JM, et al. *Biochemistry*. 7th ed. New York: W.H. Freeman; 2012.
3. Almouzni G, et al. Maintenance of Epigenetic Information. *Cold Spring Harb Perspect Biol*. 2016; 8.
4. Jenuwein T, et al. Translating the histone code. *Science*. 2001; 293: 1074-1080.
5. Forneris F, et al. Structures of C3b in complex with factors B and D give insight into complement convertase formation. *Science*. 2010; 330: 1816-1820.
6. Peuscher MH, et al. Posttranslational control of telomere maintenance and the telomere damage response. *Cell Cycle*. 2012; 11: 1524-1534.
7. Fukagawa T. Critical histone post-translational modifications for centromere function and propagation. *Cell Cycle*. 2017; 16: 1259-1265.
8. Srivastava S, et al. Posttranslational mechanisms controlling centromere function and assembly. *Curr Opin Cell Biol*. 2018; 52: 126-135.
9. Aumailley M, et al. Structure and biological activity of the extracellular matrix. *J Mol Med (Berl)*. 1998; 76: 253-265.
10. Liu Z, et al. Advantages of proteins being disordered. *Protein Sci*. 2014; 23: 539-550.
11. Marouco D, et al. Lysine-specific modifications of p53: a matter of life and death? *Oncotarget*. 2013; 4: 1556-1571.
12. McDonough AA, et al. Considerations when quantitating protein abundance by immunoblot. *Am J Physiol Cell Physiol*. 2015; 308: C426-433.
13. Cox J, et al. Quantitative, high-resolution proteomics for data-driven systems biology. *Annu Rev Biochem*. 2011; 80: 273-299.
14. Sheehan D. *Physical biochemistry : principles and applications*. 2009 [cited; Available from:
15. Kostyukevich Y, et al. Hydrogen/deuterium exchange in mass spectrometry. *Mass Spectrom Rev*. 2018.
16. Method of the Year 2015. *Nat Meth*. 2016; 13: 1.
17. Bai XC, et al. How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci*. 2015; 40: 49-57.
18. Wilson MD, et al. Cryo-electron microscopy of chromatin biology. *Acta Crystallogr D Struct Biol*. 2017; 73: 541-548.

Malattie genetiche congenite da alterazioni dell'imprinting

Orsetta Zuffardi

Dipartimento di Medicina Molecolare, Università di Pavia

Un allele la cui è espressione è sottoregulata per via epigenetica dal genitore di origine è definito come allele imprinted. I disturbi dell'imprinting sono causati da fattori genetici e variazioni epigenetiche che alterano il dosaggio di geni la cui espressione dipende appunto dal genitore di origine (geni imprinted). Più specificamente, per variazioni epigenetiche si intendono quelle che non modificano

Tab. 1 - disordini da imprinting.

Disordine	Cromosoma	Prevalenza	N° OMIM	% errore genetico (SNV/CNV)	% errore cromosomico (UPD)	% errore dell'imprinting (% MLID)	Referenze
Sindrome di Angelman	15q11.2	1:15.000	#105830	- 70% CNV (del15mat) - 15% SNV (UBE3A)	<5% (upd15pat)	<5% (rare)	Buiting, 2010
Sindrome di Prader-Willi (PWS)	15q11.2	1:15.000	#176270	70% CNV (del15pat)	<30% (upd15mat)	1%	Buiting, 2010
Sindrome di Beckwith-Wiedemann (BWS)	11p15.5	1:10.500	#130650	- 5% SNV (CDKN1C) <5% CNV and SNV of H19/IGF2 IG-DMR)	20%	10% H19/IGF2 IG-DMR hypermethylation (rare) 60% KCNQ1OT1 TSS-DMR hypomethylation (30%)	Choufani et al., 2010
Sindrome di Silver-Russell (SRS)	- 11p15.5, - chr7	1:50.000 (?)	#180860	<1%	- 10% (upd7mat) <1% (upd11mat: 11p15 LOM)	40% (15-38%)	Eggermann 2010; Wakeling et al., 2016
Pseudoipoparatiroidismo di tipo 1b (PHP1b)	20q13.3	??	#603233	27% CNV (delSTX16mat) 3% SNV (GNAS)	5% (upd20mat)	61% (rare)	Mantovani et al., 2016; Elli et al., 2016

Disordine	Cromosoma	Prevalenza	N° OMIM	% errore genetico (SNV/CNV)	% errore cromosomico (UPD)	% errore dell'imprinting (% MLID)	Referenze
Diabete mellito transitorio neonatale di tipo 1 (TNDM)	6q24	1:300.000	#601410	40% CNV (dup6pat)	40% (upd6pat)	20% (50%)	Mackay and Temple, 2010
Sindrome di Kagami-Ogata (KOS)	14q32	??	#608140	15% CNV (del14mat)	65% (upd14pat)	20% (nk)	Ogata and Kagami, 2016 Kagami et al., 2017
Sindrome di Temple (TS14)	14q32	??	#616222	10% CNV (del14pat)	78% (upd14mat)	12% (rare)	Ioannides et al., 2014; Kagami et al., 2017
Sindrome di Mulchandani-Bhoj-Conlin (MBCS)	20	??	#617352	??	100% (upd20mat)	??	Mulchandani et al., 2015
Sindrome di Schaaf-Yang (SHFYNG)	15	??	#615547	100% inattivazione di MAGEL2 (SNV/CNV)			Fountain et al., 2017
Pubertà precoce centrale di tipo 2 (CPPB2)	15	??	#615436	Inattivazione di MKRN3 (SNV)			Abreu et al., 2013

SNV: single nucleotide variant; CNV: copy number variant; upd: uniparental disomy; MLID: Multi-locus imprinting disorder

la sequenza del gene ma ne alterano l'espressione e, di conseguenza, il fenotipo associato a quel gene. I loci imprinted contengono sia geni che codificano per proteine che geni codificano per RNA che non vengono trascritti. In tutti e due i casi, la loro regolazione è mediata da elementi regolatori denominati ICR (imprinting control region) che controllano l'imprinting di uno o più geni sotto il loro controllo, limitandone l'espressione.

Tuttavia, questo processo regolatorio è molto complesso: per esempio, l'ICR sul cromosoma 14q32, sebbene metilato sul cromosoma paterno, controlla la metilazione di geni localizzati sia sul cromosoma paterno sia su quello materno. Inoltre, i geni imprinted possono essere regolati da altre sequenze di DNA oltre che dagli ICR, quindi la loro espressione può non essere imprinted dovunque: per esempio, il gene della sindrome di Angelman, UBE3A, mostra espressione imprinted solo nel cervello e il gene GNAS è espresso dall'allele materno solo in alcuni tessuti come tubuli renali e ghiandola pituitaria, ciò che spiega le caratteristiche cliniche dello Pseudoipparatiroidismo di tipo 1b. In altre parole, i geni im-

printed anche se, come regola generale, sono espressi da un solo allele parentale, in alcuni tessuti possono mantenere un'espressione biparentale. Ne consegue che le eventuali alterazioni genetiche/epigenetiche possono specificamente danneggiare quegli specifici tessuti.

Nell'insieme, i disturbi dell'imprinting (ID: imprinting disorders) hanno caratteristiche che possono rendere difficile la diagnosi per la loro eterogeneità clinica e molecolare, per la sovrapposizione di alcune caratteristiche cliniche condivise fra ID diversi, per la non rara associazione con mosaicismi somatici e, talvolta, per il coinvolgimento multi-locus delle alterazioni dell'espressione.

Nell'uomo, le regioni genomiche che mostrano imprinting, cioè metilazione del DNA differenziale (DMR) secondo il loro genitore di origine, sono circa 40 ma è possibile che ve ne siano di ulteriori non ancora identificate. Tuttavia i disordini da imprinting ad oggi noti sono solo 9 in quanto alcune regioni imprinted non sono state finora associate a specifici disordini.

Questi disordini, che come si è detto sono causati da un dosaggio alterato dei geni imprinted, possono originarsi per cause diverse quali cambiamenti genetici di tipo SNV (varianti di singoli nucleotidi), CNV (varianti del numero di copie) disomia uniparentale (UPD) e epimutazioni cioè cambiamenti della metilazione del DNA a livello delle DMR (differentially methylated regions) che controllano i geni imprinted. Inoltre, negli ultimi anni è emerso che pazienti con mutazioni epigenetiche ad un certo locus mostrano di fatto disordini della metilazione in numerosi altri loci imprinted (MLID: Multi-locus imprinting disorder), in genere in una situazione di mosaicismo.

Nella tabella I, modificata dalla pubblicazione di Mackay e Temple del 2017, sono riportati i disordini da imprinting assieme alla loro frequenza ed alle caratteristiche molecolari.

Bibliografia

1. Abreu AP, Dauber A, Maced DB, Noel SD, Brito VN, Gill JC, et al. Central precocious puberty caused by mutations in the imprinted gene MKRN3. *N. Engl. J. Med.* 2013; 368: 2467e2475.
2. Buiting K. Prader Willi syndrome and Angelman syndrome. *Am. J. Med. Genet.* 2010; 154C: 365e376.
3. Choufani S, Shuman C, Weksberg R. Beckwith-Wiedemann syndrome. *Am. J. Med. Genet.* 2010; 154C: 343e354.
4. Elli FM, Linglart A, Garin I, de Sanctis L, Bordogna P, Grybek V, et al. The prevalence of GNAS deficiency-related diseases in a large cohort of patients. *Characterized.* 2016.
5. Eggermann T, Russell/-Silver syndrome. *Am. J. Med. Genet.* 2010; 154C: 355e364.
6. Funtain MD, Aten E, Cho MT, Juusola J, Walkiewicz MA, Ray JW, et al. The phenotypic spectrum of Schaaf-Yang syndrome: 18 new affected individuals from 14 families. *Genet Med.* 2017; 19: 45e52.
7. Ioannides Y, Lokulo-Sodipe K, Mackay DJ, Davies JH, Temple IK. Temple

- syndrome: improving the recognition of an underdiagnosed chromosome 14 imprinting disorder: an analysis of 51 published cases. *J. Med. Genet.* 2014.
8. Kagami M, Mizuno S, Matsubara K, Nakabayashi K, Sano S, Fuke T, et al. Epimutations of the IG-DMR and the MEG3-DMR at the 14q32.2 imprinted region in two patients with Silver-Russell syndrome-compatible phenotype. *Eur. J. Hum. Genet.* 2015; 23: 1062e1067.
 9. Kagami M, Matsubara K, Nakabayashi K, Nakamura A, Sano S, Okamura K, et al., Genome-wide multilocus imprinting disturbance analysis in Temple syndrome and Kagami-Ogata syndrome. *Genet. Med.* 2017; 19: 476.
 10. Mackay DJ, Temple IK, Transient neonatal diabetes mellitus type 1. *Am. J. Med. Genet. C Semin. Med. Genet.* 2010; 154C: 335e342.
 11. Mackay DJ, Temple IK. Human imprinting disorders: Principles, practice, problems and progress. *Eur. J. Hum. Genet.* 2017; 60: 618-626.
 12. Mantovani G, Spada A, Elli FM. Pseudohypoparathyroidism and Gsa-cAMP-linked disorders: current view and open issues. *Nat. Rev. Endo.* 2016; 12, 347e356.
 13. Mulchandani S, Bhoj EJ, Luo M, Powell-Hamilton N, Jenny K, Gripp KW, et al. Maternal uniparental disomy of chromosome 20: a novel imprinting disorder of growth failure. *Genet. Med.* 2015.
 14. Ogata T, Kagami M. Kagami-Ogata syndrome: a clinically recognizable upd (14) pat and related disorder affecting the chromosome 14q32.2 imprinted region. *J. Hum. Genet.* 2016; 61: 87e94.
 15. Wakeling EL, Brioude F, Lokulo-Sodipe O, O'Connell M., Salem J, Blik J, et al. Diagnosis and management of Silver-Russell syndrome: first international consensus statement. *Nat. Rev. Endocrinol.* 2016.

Gene therapy returns to centre stage

Luigi Naldini

San Raffaele Telethon Institute for Gene Therapy (TIGET), San Raffaele Scientific Institute, Milan
Vita Salute San Raffaele University, Milan

Testo tratto da Nature. 2015; 526: 351-60

Gene therapy has long fascinated scientists, clinicians and the general public because of its potential to treat a disease at its genetic roots. This is achieved by counteracting or replacing a malfunctioning gene within the cells adversely affected by the condition. As simple as the concept sounds, the hurdles to put it into practice are daunting. Gene transfer must overcome complex cellular and tissue barriers to deliver new genetic information into the target cell to drive proficient expression of a therapeutic molecule without disrupting essential regulatory mechanisms. The gene-corrected cells must be present in large enough quantities to reverse the condition, escape immunological recognition and survive in the long term, or be able to transmit the modification to their progeny to sustain the benefit. Several gene-therapy trials have been performed in the past two decades for inherited diseases, cancer and chronic infections, but only a few reported clear clinical benefits and in some, individuals experienced severe adverse events related to the vectors. Overall, concern and scepticism rose over the further deployment of these strategies. But these attitudes are radically changing. A number of phase I/II gene-therapy clinical trials have reported remarkable evidence of efficacy and safety for the treatment of various severe inherited diseases of the blood, immune and nervous systems, including primary immunodeficiencies, leukodystrophies, thalassaemia, haemophilia and retinal dystrophy, as well as cancers such as B-cell malignancies (Table 1). All of these trials exploit improved vector technologies to deliver therapeutic genes. In some trials, genetic material is transferred into haematopoietic stem cells (HSCs) or T lymphocytes (T cells) *ex vivo*, and in others hepatocytes in the liver or photoreceptors in the retina are targeted directly *in vivo*. Here I review the most relevant clinical results, highlight progress in gene-transfer technologies and in our understanding of the biological processes that underpin these advances and discuss the challenges and outlook for gene therapy.

Haematopoietic-stem-cell gene therapy

HSCs have long been a preferred target for *ex vivo* gene therapy (1). Genetic modification of self-maintaining multipotent HSCs would ensure a steady supply of their gene-corrected progeny in the body. These cells have the potential to treat

conditions that manifest when mature haematopoietic lineages fail to develop or to function correctly. Given the self-renewing nature of HSCs and the need to ensure that genetic modifications are passed on to their progeny, gene correction must be stably introduced into cellular chromatin, either by vector-mediated transgene insertion or by *in situ* gene editing.

Past trials of HSC gene therapy report clear benefits to people with selected conditions, which proved the therapeutic potential of the strategy (2-10). However, they also highlight the limitations and risks of using early generation vectors that are based on gammaretroviruses (γ -RVs) (11-15). For example, these techniques offered only a limited ability to transfer genes into the most primitive progenitor cells, giving rise to the low-level and transient appearance of gene-corrected haematopoietic cells *in vivo*. Leukaemia related to vector insertion near oncogenes also occurred in a fraction of patients during their long-term follow-up. The development of vectors with improved efficacy and safety in preclinical models, such as lentiviral vectors, has renewed interest in the approach (1).

HSC gene therapies that incorporate these new vectors have been tested in severe inherited diseases of the immune system (Wiskott-Aldrich syndrome (WAS) and X-linked severe combined immunodeficiency (SCID-X1)) (8,16-18) and blood (β -thalassaemia) (19), and in neurodegenerative storage diseases (adrenoleukodystrophy and metachromatic leukodystrophy) (20-22) (Table 1). In children with SCID-X1, the development and function of the immune system is impaired owing to deficiencies in the receptors for certain cytokines that are essential for immune-cell development, whereas those with WAS have deficiencies in a cytoskeletal adaptor that is required for assembling the immunological synapse (23). Consequently, these children succumb to infection or, in the case of WAS, haemorrhage because of an accompanying platelet deficiency. People with β -thalassaemia major - the most severe form of β -thalassaemia - fail to express the haemoglobin- β chain, which leads to ineffective erythropoiesis and severe anaemia that requires frequent blood transfusions and an iron-chelation regimen. Children with early onset adrenoleukodystrophy or metachromatic leukodystrophy are affected by defective myelination and consequently experience glial- and neural-cell degeneration in the central and peripheral nervous systems. They are unable to break down some of the metabolites of myelin because of a deficiency in the peroxisomal ATP-binding cassette transporter in adrenoleukodystrophy (24), or in the lysosomal enzyme arylsulfatase A in metachromatic leukodystrophy (25). Such patients undergo rapid and irreversible deterioration of their motor, sensory and cognitive performance, which leads to death within a few years. HSC transplantation, in which HSCs are transferred from a healthy donor to a recipient with a specific condition, can virtually cure SCID-X1, WAS and β -thalassaemia (26). HSC transplantation can also arrest the progression of leukodystrophies if performed before or near the time of onset, although the outcome of the procedure is more satisfactory in adrenoleukodystrophy (24) than in metachromatic leukodystrophy (27). Allogeneic HSC transplantation confers a considerable risk of morbidity and death, particularly when performed between human leukocyte antigen (HLA)-mismatched individuals. Even in the few cases in which an HLA-matched

Table 1 | Gene-therapy clinical trials highlighted in this Review

Disease	Vector and strategy	Number of patients*	Follow-up (months)	Patient status and biological and clinical outcomes*	Clinical-trial identifier	References
HSC-based gene therapy						
Wiskott–Aldrich syndrome	Lentiviral vector; <i>ex vivo</i> gene transfer into CD34 ⁺ cells.	7	10 to 60	All patients AAW; stable engraftment with transduced cells; persistent clinical benefit and safety.	NCT01515462	16 and L.N.†
Wiskott–Aldrich syndrome	Lentiviral vector; <i>ex vivo</i> gene transfer into CD34 ⁺ cells.	7	9 to 42	6 patients AAW, 1 patient died of a pre-existing infection; stable engraftment with transduced cells; persistent clinical benefit and safety.	NCT01347242 NCT01347346 NCT02333760	17
X-linked severe combined immunodeficiency	Self-inactivating γ -RV; <i>ex vivo</i> gene transfer into CD34 ⁺ cells.	9	12 to 39	8 patients AAW, 1 patient died of an infection; stable engraftment with transduced cells; persistent clinical benefit and safety in 7 patients; 1 patient failed to engraft and underwent HSC transplantation.	NCT01410019 NCT01175239 NCT01129544	18
β -Thalassaemia major	Lentiviral vector; <i>ex vivo</i> gene transfer into CD34 ⁺ cells.	3‡	24 to 72	2 patients stably engrafted with transduced cells, 1 patient became transfusion independent; 1 patient failed to engraft and received rescue cells.	N/A	19, M. Cavazzana and BlueBird Bio†
β -Thalassaemia major	Lentiviral vector; <i>ex vivo</i> gene transfer into CD34 ⁺ cells.	2‡	15	Stable engraftment with transduced cells; transfusion independence and safety in both patients.	NCT02151526	M. Cavazzana and BlueBird Bio†
β -Thalassaemia major	Lentiviral vector; <i>ex vivo</i> gene transfer into CD34 ⁺ cells.	5§	1 to 6	Stable engraftment with transduced cells; safety and transfusion independence in the first 2 evaluable patients.	NCT01745120	BlueBird Bio†
Adrenoleukodystrophy	Lentiviral vector; <i>ex vivo</i> gene transfer into CD34 ⁺ cells.	4	54 to 101	Stable engraftment with transduced cells and safety in all patients; persistent clinical benefit in 3 patients.	N/A	20, 21 and P. Aubourg†
Metachromatic leukodystrophy	Lentiviral vector; <i>ex vivo</i> gene transfer into CD34 ⁺ cells.	20	3 to 60	Stable engraftment with transduced cells and safety in all patients; persistent clinical benefit in all late-infantile patients who were treated when presymptomatic.	NCT01560182	22 and L.N.†
Liver-directed gene therapy						
Haemophilia B	AAV8 vector; intravenous administration.	10	16 to 48	No inhibitors; persistent FIX expression; in high-dose group, mean FIX levels of $5.1 \pm 1.7\%$ seen in all 6 treated patients.	NCT00979238	40
Haemophilia B	AAV8 vector; intravenous administration.	7	Up to 12	No inhibitors; persistent FIX expression in 1 patient.	NCT01687608	110
T-cell immunotherapy for cancer						
B-cell lymphoma or CLL	γ -RV; <i>ex vivo</i> gene transfer into T cells; CAR-modified anti-CD19 cells.	15	1 to 23	8 CRs, 4 PRs; ORR 80%.	NCT00924326	51
B-cell ALL	γ -RV; <i>ex vivo</i> gene transfer into T cells; CAR-modified anti-CD19 cells.	5	1 to 4	5 CRs; ORR 100%; 4 patients subsequently underwent allo-HSC transplantation as per clinical-study design, 1 patient was ineligible for HSC transplantation and relapsed.	NCT01044069	52
B-cell ALL	Lentiviral vector; <i>ex vivo</i> gene transfer into T cells; CAR-modified anti-CD19 cells.	30	1 to 24¶	27 CRs; ORR 90%; 19 patients remained in remission, 3 of these patients underwent HSC transplantation; 7 patients relapsed, 3 of these relapses occurred after loss of transduced T cells.	NCT01626495 NCT01029366	53
B-cell ALL or lymphoma	γ -RV; <i>ex vivo</i> gene transfer into T cells; CAR-modified anti-CD19 cells.	21	Median = 10#	14 CRs (at day 28); ORR 67%; 10 patients subsequently underwent HSC transplantation.	NCT01593696	54
Retinal gene therapy						
Type 2 Leber congenital amaurosis	AAV2 vector; unilateral subretinal administration.	5	36	In all 5 patients, stable improvement in visual sensitivity seen.	NCT00516477	94, 97
Type 2 Leber congenital amaurosis	AAV2 vector; unilateral subretinal administration.	3	54 to 72	In all 3 patients, improvement in visual sensitivity seen at 6 months, which increased for 1 to 3 years and then declined.	NCT00481546	95
Type 2 Leber congenital amaurosis	AAV2 vector; unilateral subretinal administration.	12	36	In 6 patients, improvement in visual sensitivity seen, which peaked at 6 to 12 months and then declined.	NCT00643747	96

AAW, alive and well; ALL, acute lymphocytic leukaemia; CLL, chronic lymphocytic leukaemia; CR, complete response; N/A, not applicable; ORR, overall response rate; PR, partial response.

*As stated in the referenced publications or updated by personal communication. †Personal communication. ‡ β^0/β^0 genotype. § $2\beta^0/\beta^0$, $2\beta^0/\beta^0$, $1\beta^0/\beta^0$ genotypes. ¶Median = 7, #51 days to HSC transplantation median.

family donor can be found, the risk of morbidity remains substantial (28). HSC gene therapy can address this unmet medical need, especially when no matched HSC donor is available.

HSC gene therapy is administered by *ex vivo* gene transfer into haematopoietic progenitors. First, the cells are purified using the CD34 surface marker from other leukocytes harvested from the bone marrow or mobilized peripheral blood of the recipient. Next, they are cultured for 2-4 days in the presence of growth-stimulating cytokines while being exposed to vectors carrying an expression cassette for the corrective transgene. Before the modified cells can be administered, the recipient is treated with a preconditioning chemotherapy regimen. This depletes endogenous progenitors and differentiated cells in the bone marrow - as well as the lymphoid organs, in some cases - and favours engraftment of the *ex vivo* gene-corrected cells. Preconditioning results in considerable early morbidity owing to transitory blood-cell depletion, immunodeficiency and mucosal damage, which place the recipient at risk of severe infection. It also causes delayed morbidity owing to the risk of developing chemotherapy-induced secondary tumours and infertility (28). Several HSC gene-therapy trials have attempted to alleviate the morbidity associated with preconditioning by lowering the dosage and combination of chemotherapeutic drugs that are administered in comparison to the regimens used in conventional HSC transplantation. However, the impact of changing these drug regimens on the risks and benefits of the therapy is yet to be determined in broader comparative studies and through long-term patient follow-up. Some preconditioning regimens also deplete endogenous microglia progenitors, which live in the central nervous system (CNS). In patients with leukodystrophies who have been treated with HSC gene therapy, these cells are replaced with their gene-corrected counterparts, which migrate to CNS tissues and produce functional progeny that can clear stored myelin metabolites. In the case of metachromatic leukodystrophy, these modified cells also release functional lysosomal enzyme for the cross-correction of other tissue-resident cells (29-31).

All HSC gene-therapy trials performed with lentiviral vectors report stable and high-level reconstitution of haematopoiesis with gene-corrected cells in most recipients, with the most recent trials observing up to 90% reconstitution in some patients (16, 17, 19-22). Because all of the haematopoietic lineages tested (myelo-monocytic, megakaryocytic, erythroid, natural killer (NK), and B lymphoid and T lymphoid) contain gene-corrected cells and appear to receive a steady input of newly formed modified cells - most evident in the short-lived myeloid lineages - gene transfer must have occurred in self-renewing, multipotent progenitors that effectively engrafted to the recipients. The extent of haematopoietic reconstitution with gene-corrected cells varies among the patients and trials. This reflects the impact of the disease and its correction on the different haematopoietic lineages as well as the potency of each vector batch and the type of preconditioning regimen administered. Remarkably, most patients gain substantial benefits from gene therapy that can even exceed those observed after successful HSC transplantation from an allogeneic donor. This is especially true for a metachromatic

leukodystrophy trial (22) based on the rationale that gene transfer could drive higher lysosomal enzyme expression in haematopoietic cells than in normal cells, which would enable these cells to provide an increased supply of enzyme while homing in to affected tissues. The strategy allowed the unsatisfactory outcome of HSC transplantation to be overcome in metachromatic leukodystrophy. Despite an increasing number of patients being treated in these HSC gene-therapy trials (Table 1), there has been no report of adverse events related to lentiviral vectors. At the most recent follow-up of patients included in these trials, the first of whom was treated almost seven years ago in the adrenoleukodystrophy trial (20), those who had successfully engrafted with modified cells were reported to be in a good condition - that is, they are either disease free or their disease has been stable or in remission since the time of treatment - and are leading near-normal lives. These individuals would probably have already succumbed to their disease if untreated with gene therapy (24, 25).

Liver-directed gene therapy

The liver has long been a preferred target for *in vivo* gene therapy (32). This major internal organ and central metabolic hub receives an abundant blood supply through an extensive bed of sinusoids with highly permeable walls - a structure that facilitates the access of blood-borne particles, such as viruses, to hepatocytes. Hepatocytes are long-lived and robust protein factories that can efficiently release their products into the blood circulation. Stable transgene delivery to the liver could therefore provide a strategy for treating several inherited metabolic diseases and plasma-protein deficiencies, notably those of coagulation factors. Major hurdles to liver-directed gene therapy include the potential toxicity of an acute inflammatory response to the bolus administration of viral particles into the bloodstream, and the inactivation of these particles by pre-existing virus-specific antibodies and clearance by phagocytes that line the sinusoidal walls of the liver and spleen. If humoral or cellular immunity is triggered against the transgene product, the therapeutic activity of gene therapy could be inhibited in the circulation or the modified cells might be eliminated by cytotoxic T cells, respectively (33, 34).

Liver-directed gene therapy has been tested mainly in the treatment of severe haemophilia B using vectors derived from the human parvovirus, adeno-associated virus (AAV) (35). On intravascular injection, some AAV-vector serotypes effectively target hepatocytes and remain stably associated with the modified cells, mostly as nonintegrated episomes within the nucleus. In animal models, this delivery strategy does not induce an immune response against the transgene product and instead favours the development of transgene-specific tolerance (32). The persistence of AAV-vector genomes that contain an expression cassette for coagulation factor IX (FIX) - a protein that is absent or defective in haemophilia B - within nonproliferating hepatocytes can give rise to a stable supply of functional FIX into the bloodstream (36). FIX expression as low as 1% of the normal level can turn severe haemophilia B into a milder form of disease. This partial reconstitution alleviates the risk of spontaneous haemor-

rhages, which are detrimental to the joints and potentially lethal if they occur in the brain. It can also reduce the need for intravenous prophylactic factor-replacement therapy, a treatment that imposes a substantial burden on patients and is only available at high cost and in countries with well-developed health-care systems (37). Although earlier trials demonstrated the safety of delivering AAV vectors into the human bloodstream, they reported only limited and transient FIX expression. This is probably because of low levels of gene transduction in the liver and a delayed cellular immune response that targeted AAV components persisting within modified hepatocytes and triggered their elimination (38, 39). More recently, a trial that used a new AAV serotype (AAV8) and an improved cassette design was able to overcome these limitations to demonstrate vector-dose-dependent FIX expression of up to 6% of the normal level after a single, well-tolerated vector infusion in a peripheral vein (40). Three years later, FIX expression remained stable in most patients, although some had to be given a transient immunosuppressive corticosteroid treatment at the first sign of hepatocellular injury in the 4-8 weeks following treatment. Most treated patients were able to reduce or abrogate the need for prophylactic factor replacement, which substantially improved their quality of life.

T-cell immunotherapy for cancer

T cells are also popular targets for *ex vivo* gene therapy. Such therapies aim mostly at boosting the adaptive immune response against cancer and chronic infections such as HIV (41, 42). Autologous T cells can be harvested readily from the peripheral blood and expanded *ex vivo*. Cells are then transduced with a -RV or lentiviral vector expressing an exogenous T-cell antigen receptor (TCR) that is specific to a cancer-associated antigen or an antiviral molecule and infused back into the patient. The use of T cells for cancer immunotherapy arose from seminal observations of objective tumour responses, which sometimes led to complete tumour regression, after the infusion of *ex vivo*-expanded autologous tumour-infiltrating lymphocytes in patients with advanced melanoma (41). The underlying mechanisms of these responses have been traced to the *ex vivo* activation and amplification - and subsequent *in vivo* persistence - of tumour-specific cytotoxic T cells that already exist in small numbers within some tumours but are suppressed by the local microenvironment. These data support the therapeutic potential of an adaptive immune response against some tumours when it is released from endogenous suppressive signals and the reactive T cells are infused at high numbers, which leads to a favourable effector-target ratio. Preconditioning patients with a lymphoid-depleting regimen before infusion of the *ex vivo*-expanded T cells improves their engraftment and activity *in vivo* (43).

The genetic transfer of an exogenous TCR to a T cell can bypass the need to find preexisting cancer specificities within the patient's T-cell population. It also allows the exploitation of synthetic high-affinity TCRs, which might otherwise be purged *in vivo* by thymic selection, to redirect the specificity of autologous cells. Ideally, such TCRs would target tumour-associated antigens from endogenous proteins that are not being expressed - at the same time or

at detectable levels - by normal tissues, such as carcinoembryonic or cancer testis antigens. Alternatively, they could target tumour-driver mutations that are commonly found within a certain type of tumour and that cannot be lost, even to evade immune clearance. Recent studies, however, indicate that most spontaneous or elicited tumour-specific immune reactivity is instead directed against 'passenger' neoantigens, which uniquely originate from random mutations that accumulate within individual tumours (44-47).

This type of immunotherapy must therefore be highly personalized and will require both the identification of candidate neoantigens from the tumour exome or proteome and the retrieval or generation of the cognate TCR recognition sequences for *ad hoc* T-cell genetic engineering.

Early clinical trials of T-cell gene transfer with TCRs directed against tumour-associated antigens reported partial tumour responses and, in some instances, off-tumour reactivity that led to tissue damage and severe adverse events (48-50). More recently, gene transfer was used to introduce synthetic chimaeric antigen receptors (CARs) to T cells. CARs combine the binding specificity of an antibody against a cancer-associated surface marker with one or more intracellular signalling domains from the TCR and costimulatory receptor complexes (42). The strategy holds a number of advantages over the use of conventional TCRs. For instance, antigen recognition is not restricted by HLA, and antibody specificities previously validated for safe and specific cancer recognition *in vivo* can be exploited by the incorporation of single-chain antibody derivatives into the CARs. Furthermore, these engineered T cells can be fully activated when meeting their target. Trials that deployed CARs directed against the B-cell surface molecule CD19 reported dramatic benefits in patients with B-cell malignancies, who experienced durable clinical responses, including complete remission, with mostly manageable toxicity (51-54). These results have spurred enormous interest in further developing this approach (55).

Reasons for success in recent clinical trials

The positive clinical outcomes discussed in this Review provide long-sought evidence for the elusive promise of gene therapy to deliver lasting therapeutic benefit, or even a 'cure', for an otherwise terminal or severely disabling condition after a single treatment. These results could represent the rewarding outcome of the effort to rationally improve vectors in the laboratory, as discussed in Box 1 and Box 2. Advances in vector manufacturing and characterization, leading to batches with higher potency and greater purity, could also be facilitating the increased transduction of target cells with lower adverse effects, even when using the same vector design as in earlier trials. A better understanding of the biological processes and specific cell types that are involved in the success or failure of each gene therapy has also helped to improve methods for *ex vivo* cell handling and led to more effective monitoring and management of patients. Indeed, increasing confidence in the positive outcome of recent trials is now built on both clinical observation and advanced molecular readouts that provide evidence for the efficacy and safety of gene therapy.

BOX 1 - Lentiviral vectors

Major hurdles for haematopoietic-stem-cell (HSC) gene therapy include achieving efficient *ex vivo* gene transfer into long-term repopulating hscs, preventing activation of oncogenes by the nearby integration of a vector and controlling transgene expression to avoid ectopic or constitutive expression that leads to toxicity (1). As compared to early generation gammaretroviral vectors (γ -rvs), HIV-derived lentiviral vectors result in more efficient gene transfer and stable, robust transgene expression in hscs and their multilineage progeny. Extensive preclinical work indicated important features in vector biology and design that affect genotoxicity and highlighted strategies to alleviate it (111-117). The self-inactivating long terminal repeats (ltrs) and integration-site preferences of lentiviral vectors were shown to substantially alleviate insertional genotoxicity. When tested in γ -rvs, the self-inactivating LTR design was shown to improve the safety of this platform as well (18). Retrospective analysis of several earlier trials suggests that disease background, transgene function, *ex vivo* culture and the efficiency of host repopulation can all influence the likelihood that insertional genotoxicity will manifest in a trial (13, 14). These data helped to shape the ideas that not all integrating vectors have the same effects and that genome-wide integration of improved vector designs, although still a mutagenic event, can be tolerated in the absence of aggravating circumstances (118).

Self-inactivating lentiviral vectors are also being used to engineer T cells with chimeric antigen receptors (cars) or T-cell antigen receptors for use in adoptive immunotherapy for the treatment of cancer. The advantages of this new platform in comparison to earlier-generation γ -rvs, which perform satisfactorily in this cell target, are yet to be fully established. Lentiviral vectors are thought to give rise to more robust and stable transgene expression in T cells *in vivo*, and could facilitate more efficient and versatile *ex vivo* gene transfer while supporting coordinated expression of multiple transgenes (41, 42, 55, 119). These advantages will become more relevant as the gene-therapy field implements refined strategies, such as improved T-cell manipulation to preserve T memory stem cells (59-61), or more demanding cell-engineering tasks, such as the co-expression of multiple cars (to improve specificity) or a conditional safety switch/suicide gene (to improve safety) (120).

In recent lentiviral-vector-based HSC gene-therapy trials, vector insertional analysis was used to track individual clones in the reconstituted haematopoiesis and showed that polyclonal reconstitution by transduced stem cells had taken place without the emergence of dominant clones whose behaviour could be attributed to gain-of-function vector insertional mutagenesis. An in-depth molecular analysis of the reconstituted haematopoiesis now being performed in different diseases, including several that were treated with different types of vector, is facilitating the first reliable comparative assessment of vector-induced events in patients (16, 17, 20-22, 56). The emerging picture illustrates that the absence of adverse clinical events in recent trials is accompanied by a remarkably different landscape of vector-insertion distribution in patients. In the earlier trials, which were performed with long terminal repeat (LTR)-competent γ -RV vectors, there was frequent and early generation of dominant clones, whose expansion appeared to be driven by the altered expression of cancer genes that were targeted by vector insertion (11-13, 14, 57). However, this finding is rarely observed when newer vectors, such as lentiviral vectors, with self-inactivating LTRs and different insertion-site preferences are used (16, 17, 19-22). In parallel, retrospective

BOX 2 - New adeno-associated virus vector serotypes

Major hurdles for liver-directed gene transfer include alleviating the inflammatory response to intravenous administration of high-dose viral particles and bypassing pre-existing immunity to viral components. Hepatocyte targeting must also be improved and long-term transgene expression should be established (33, 34). In a recent clinical trial of liver-directed haemophilia B gene therapy, an adeno-associated virus (AAV) vector that had shown improved liver tropism (provided by the AAV8 serotype) and enhanced transduction efficiency (obtained by packaging a self-complementary genome) in animal models, resulted in the long-term expression of factor IX (FIX) in most patients (40) as long as a transient immunosuppressive corticosteroid treatment was promptly administered to those who were developing signs of hepatocellular injury. Whether the successful outcome can be attributed to the use of the new vector design remains to be established because the dose-response relationship and tropism of the different vector serotypes might differ between humans and animal models (64). Another ongoing trial of haemophilia B gene therapy is using the same AAV8 serotype to express a hyperactive *FIX* transgene. Preliminary reports indicate that sustained FIX expression occurs only in some patients, despite the administration of oral corticosteroids on signs of hepatocellular injury (110). It is possible that the immunosuppressive treatment was administered too late to be effective at preventing the clearance of modified hepatocytes. Alternatively, certain aspects of vector design and manufacturing might be missed by current methods of characterization, which could influence the robustness of the therapy.

reconstruction of multistep leukaemogenesis from the earlier trials supports the view that vector insertion provided a first mutagenic hit that potentially led to the development of leukaemia. Progression of the disease was facilitated by accumulation of further vector-independent mutations in the expanding clone, which were favoured by concurrent precipitating factors such as oligoclonal reconstitution, selective pressure for a survival or growth advantage, transgene toxicity or stressed haematopoiesis resulting from the underlying disease (12, 14). Overall, these findings support the idea that improved efficiency of HSC transduction and transplantation, together with improved vector design, can substantially alleviate, but not eliminate, concerns about genotoxicity. Once it has been established that vector insertion is mostly neutral to cell behaviour, tracking clonal activity in the reconstituted haematopoiesis can highlight important biological features of haematopoietic regeneration. These include the pattern of activity and stability of the transduced HSCs, which provides the foundation for predicting long-term maintenance of the therapeutic benefit (16, 22). Tracking clones can also help to establish the reliability of current models of lineage relationship and haematopoietic hierarchy based on xenotransplantation studies (58).

In adoptive T-cell therapy, a subpopulation of T memory stem cells plays an important part in supporting long-term efficacy. This knowledge arose from an improved understanding and phenotypic characterization of T-cell differentiation pathways, and the identification of the cells that are responsible for the sustained generation of effector activity *in vivo* (59). Optimization of *ex vivo* cell culture (60) and *in vivo* tracking of gene-marked T cells (61) also contributed to this discovery, which should facilitate the design of more effective trials. Epitope

mapping and tracking of cancer-targeting T cells in patients who are undergoing adoptive T-cell therapies and checkpoint blockade drug treatment provide direct evidence for the potential to evoke cancer-specific effector T cells in some types of cancer, which could then be exploited to clear advanced metastatic disease after *ex vivo* amplification or genetic editing (44-47).

In liver-directed gene therapy, detailed epitope-specific immune monitoring of antiviral T-cell responses on *in vivo* AAV gene delivery is able to account partially for earlier failures to observe long-term stable FIX expression (34, 39). This also helped to uncover strategies for controlling a delayed cytotoxic adaptive response to viral components that persist within transduced cells, which include administering transient immunosuppressive drugs at the first occurrence of such a response (40) and engineering viral capsids to bypass dominant responses (62-64).

The enhanced efficacy observed in recent gene-therapy trials could also reflect improved study designs. These have been made possible by a deeper understanding of vectorhost interactions and by the application of more rational patient-selection criteria, which are more easily adopted after the initial 'proof of safety' of administration by vector in humans has been obtained. Improved trial designs can enhance efficacy in several ways. For example, treating early symptomatic (or even presymptomatic) individuals instead of people with an advanced stage of a disease provides a greater chance for the therapy to work before tissue damage has become irreversible. Vector doses can be set at levels that approach therapeutic efficacy instead of minimal biological activity, as is usually done to alleviate risks when establishing the safety of a treatment in phase I/II trials. However, this cautionary approach often denies trial participants the chance to gain immediate benefit from the therapy, as well as any future benefit (because they become immune to the vector). By administering more aggressive preconditioning regimens, space can be made for the gene-corrected cells. Vector-induced adverse events can also be neutralized by preemptive or prompt pharmacological treatment, such as administering anti-inflammatory drugs or steroids to suppress innate or adaptive immune responses, respectively.

Gene-therapy trials often address rare diseases for which little is known about the natural history and genotype-phenotype relationships. Enriching such knowledge helps in the design of more effective trials because it becomes possible to recruit the patients who are most likely to benefit from the treatment. Efficacy can also be established according to validated endpoints within a timeframe that is compatible with further development of the therapy.

Challenges ahead and prospective developments

In HSC gene therapy, despite advances in vector design, determining the actual risk of insertional mutagenesis in clinical trials - and how it can be managed at the level of the individual patient or overcome by further improvements in vector design - remain major targets for future progress. Several experimental models have been developed to assess and rank the relative genotoxicity of different vector types and designs. Despite providing the rationale to advance new vectors to

clinical testing, they fail to provide a quantitative prediction of actual oncogenic risk in the clinical setting. It is still difficult to establish the actual risk of genotoxicity for a given vector in a given trial, especially considering the low number of patients that have been treated with vectors to date, the longer follow-up periods that are required, and the possibility that the disease background could concur to increase the risk. In addition, it remains unclear whether molecular monitoring can predict the progression to malignancy of aberrant clones in individual patients. As long as the chosen vectors are able to integrate throughout the genome, subject to preferences dictated by the parental virus and particle composition, the risk of gene activation and inactivation at insertion sites can be mitigated but not abrogated. Although the improved vector types and designs that are currently being tested could alleviate some of the concerns that surround oncogene activation, which represents a high-risk dominant mutagenic event, the disruption of tumour-suppressor genes remains possible. These less frequent and recessive events might reveal their consequences only after a longer latency period or through testing in large-scale trials.

If efficacy and safety can be established in greater numbers of patients and over longer follow-up periods, HSC gene therapy might eventually challenge the dominance of allogeneic HSC transplantation as a first-line therapeutic option in monogenic diseases for which HSC transplantation is beneficial (15, 65). In fact, the use of autologous HSCs makes the treatment potentially available to all patients. HSC transplantation, meanwhile, is available only to patients for whom an HLA-matched or compatible donor can be found. HSC gene therapy can also substantially lower morbidity because it abrogates the risk of graft-versus-host disease and abolishes the need for immune suppression after treatment. Moreover, it could permit the use of milder preconditioning treatments, as partial chimaerism through the presence of gene-corrected cells might be sufficient to correct the disease and spare the patients the risk of myelo- and lymphoablative preconditioning treatments. In addition, genetic engineering of HSCs might facilitate new modes of treatment if the present outcome of HSC transplantation is unsatisfactory. This could involve increasing the therapeutic gene dosage to a level that is higher than the level that is provided by normal donor cells (as previously described for metachromatic leukodystrophy) or delivering transgene activity to selected tissues or sites of disease (such as the CNS in metachromatic leukodystrophy). HSCs and their progeny could also be equipped with exogenous genetic information to better fight cancer or chronic infection. As therapeutic options become available for otherwise incurable diseases, the genomes of newborn babies could be screened to allow early treatment of these conditions. To fulfil these predictions, *ex vivo* genetic modification and culturing of HSCs and their progenitors must become more robust and reliable. This should be coupled with the improved maintenance and expansion of the treated cells, which would support faster haematopoietic recovery in preconditioned patients and increase the clonal composition, resilience and stability of the engineered graft - thereby improving the short-term and long-term safety of the procedure. When therapeutic benefit is reached on establishing a partial chimaerism, less-toxic conditioning regimens based on biological agents

could also be applied, which would spare patients both the acute and long-term toxicity of current chemotherapy-based regimens (66).

In T-cell immunotherapy for cancer, the number of CARs in clinical testing is growing quickly, although most of the trials focus on treating B-cell malignancies using CARs that target B-cell surface antigens. Notably, this approach depletes not just the disease-causing malignant clones but also almost all B cells in the patient. Although the lack of B cells can be remedied by the infusion of immunoglobulins, depletion of other cell lineages might not be as manageable. This issue could limit the use of other lineage-specific antigens as CAR targets. In addition, the clearance of large tumour masses observed in these trials was accompanied by an acute and often severe syndrome - even requiring intensive care - that followed the massive release of cytokines from on-target activated T cells.

Given the remarkable benefits that have been observed in some patients, further CARs are likely to be designed and tested (42, 55). However, important questions still need to be addressed. For example, what CAR design will achieve the greatest efficacy with minimal toxicity? How can the toxicities deriving from the CAR-mediated recognition of healthy tissues that express target antigens at low levels be controlled in a timely and effective way? Can adoptive T-cell therapies be as effective in solid cancers as they seem to be in some lymphoid malignancies, or might features of the tumour stroma inhibit their activity? Although more challenging than transferring CAR genes, transferring TCR genes could be more effective in the long term. TCR gene transfer could also be a better fit for cancers that tend not to accumulate passenger mutations and might be less responsive to checkpoint blockade drugs. By targetting cancer-driver mutations, TCR gene transfer might be less prone to the induction of resistance than CAR-gene transfer, which has a more limited range of targets that must be surface antigens and are not always drivers of the disease. TCR gene editing (67, 68) can redirect specificity of a T cell towards a new antigen by disrupting endogenous TCR genes then introducing an exogenous TCR or CAR. It could therefore become a powerful strategy to avoid TCR mispairing, which occurs when the same cell expresses two different TCRs. Moreover, in driving the biological responses of the genetically modified T cells, it suppresses confounding endogenous TCR signal transduction. When combined with growing evidence to support the substantial efficacy of immune checkpoint blockade therapy, the outcome of recent CAR T-cell-gene transfer trials suggests that immunotherapy could become a new pillar of cancer therapy that has the potential to eradicate the disease (69).

In some of the liver-directed, AAV-based, gene-therapy trials discussed in this Review, the consequences of a delayed cell-mediated immune response that targets transduced hepatocytes could be controlled. However, this remains an area of close scrutiny because the factors that might concur to trigger the response and regulate its timing are not fully understood (34). For instance, how does the occurrence of this delayed hepatocellular toxicity relate to the administered vector dose, potency and type? And is it related to possible concomitant triggers such as transient subclinical liver injury, or inflammation that recruits memory T cells at a time when viral antigen is still being presented by the transduced hepatocytes?

Another unresolved aspect concerns the durability of transgene expression over an extended period of time. Although AAV-mediated FIX expression remains stable over the life of treated dogs, it is difficult to predict whether this also applies to humans, who have longer lifespans. A clear understanding of the molecular forms that underlie the long-term association of AAV DNA with the hepatocyte nucleus - and their relative contribution to transgene expression - will surely provide insight. In preclinical studies, a fraction of the AAV genome is reported to have integrated into the chromatin of hepatocytes, although it is unclear how much it could contribute to long-term gene expression (35). Although integration ensures long-term expression of genes, even in the face of cell replication, there are concerns over its safety (70-72). Approaches must be developed that enable patients with high concentrations of AAV-neutralizing antibodies to access AAV-mediated gene therapy and to allow re-administration of gene therapy, possibly by exploiting alternative capsid composition (34, 63, 73, 74).

As confidence grows in the choice of vector type and dosing, patient selection criteria can be used to target therapies to individuals who are less likely to mount AAV-directed immune responses. New vector serotypes will ensure that AAV-based gene therapy becomes more widely applicable. For instance, it should be possible to tackle haemophilia A - the most common and challenging form of haemophilia - as well as several other metabolic and storage diseases that affect the liver and other peripheral organs, including muscle and the heart. The remarkable ability of AAV gene therapy to establish a long-term clinical benefit through a simple, well-tolerated intravenous infusion, combined with the ease of manufacturing a vector-based medicinal product (instead of a personalized cell product that is necessary for *ex vivo* gene therapy), could set the stage for its rapid commercial development and broad market distribution. AAV gene therapy might also be suitable for addressing unmet medical needs in countries with less well-developed health-care systems because a single intervention could alleviate the burden of providing and accessing lifelong replacement therapy and medical care. Other vector platforms, including lentiviral vectors, are also being developed for liver-directed gene therapy (75) and could eventually complement AAV-based delivery to broaden the access of patients to - and the diseases targeted by - this promising type of gene therapy.

Targeted gene editing

Another important boost to the gene therapy renaissance has emerged from advances in gene-targeting technologies (76). The ability to generate artificial DNA endonucleases that bind specifically to a DNA sequence of choice and induce a double-strand break (DSB) is making targeted genome editing more efficient and much easier to undertake, which brings the long-sought goals of somatic gene disruption, targeted transgene integration and *in situ* gene correction within the reach of gene therapy. All of these strategies entail a 'hit-and-run' mechanism that requires only transient expression of the nuclease complex and, in some cases, a repair template to modify the genome permanently. Because they target a selected region of the genome, these strategies abrogate the risk of insertional mutagenesis

and poorly controlled transgene expression that is associated with conventional gene-replacement strategies.

Gene targeting has recently entered clinical testing through the adoptive T-cell therapy of patients infected with HIV (77). Artificial zinc-finger nucleases (ZFNs) were used to disrupt the gene that encodes CCR5, a cellular co-receptor for HIV, in T cells grown *ex vivo* from each patient - with the aim of making these cells resistant to infection with the virus before their reinfusion. Gene disruption was achieved by creating a DNA DSB in an exon of the *CCR5* gene because repair of DSBs by the non-homologous end-joining (NHEJ) pathway leads to the loss or insertion of bases during end rejoining, which inactivates the coding sequence. The trial reported both proof of safety and long-term persistence of the engineered cells *in vivo*. It also provided an indication of efficacy through a trend for the positive selection of cells with disrupted *CCR5* alleles, as well as an observation in some patients of improved control of viral replication on the scheduled interruption of antiviral therapy (77). To achieve definitive viral clearance, however, the fraction and long-term maintenance of infused cells bearing biallelic *CCR5* disruption might need to be increased. This could be achieved by performing gene editing in self-renewing progenitor cells, such as HSCs (78) or T memory stem cells, because these cells offer better potential for achieving this level of enhanced and long-term reconstitution of the T-cell compartment, especially if combined with preconditioning to deplete endogenous non-modified cells. Targeted gene disruption can also be used to relieve the repression of an endogenous gene whose expression might compensate for the dysfunction of another gene. For instance, inactivation of the transcriptional repressor *BCL11A* in erythroid progenitor cells could reactivate fetal γ -globin expression to compensate for or counteract a dysfunctional β -globin chain in β -thalassaemia or sickle-cell disease, respectively (79). However, a higher efficiency of biallelic knockout is required to fulfil these goals when there is no mechanism by which to amplify selectively the genetically modified cells.

Gene editing becomes even more ambitious when aiming to replace a target sequence with an exogenous version of choice by exploiting the homology-directed repair (HDR) pathway of DNA DSBs. To achieve this, artificial nucleases and an exogenous DNA template bearing homology to the target site and comprising the new sequence (76) must be delivered to the cell. The approach has great potential for use in *ex vivo* gene therapy because the targeted integration of an expression cassette into a preselected genomic 'safe harbour' (80, 81) or the *in situ* reconstitution of a mutant gene would ensure robust and predictable expression - closely recapitulating the endogenous expression control in the latter - without the risk of insertional mutagenesis (82-84). Several hurdles must be overcome before these strategies can be fully exploited. This is because the efficiency of HDR-mediated genome editing remains low in most primary cell types of relevance to gene therapy, such as HSCs (82). In addition, it is challenging to achieve the safe and feasible clinical translation of cell-therapy products when having to rely on selection and extensive *ex vivo* amplification of a few edited cell clones. The cellular response to DNA DSBs varies according to cell type and cell cycle and growth

statuses, and ranges from repair by the different pathways to differentiation or apoptosis (85). Overall, how the cell chooses between NHEJ and HDR is poorly understood. Alternative mechanisms for HDR that might function outside the S/G₂-phases of the cell cycle are also emerging.

Multiplexed, targeted genome editing is now easily achievable through the application of clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 RNA-based nucleases. These nucleases can be rapidly and easily adapted to seek out any DNA target site by designing an RNA guide instead of generating a protein-based sequence recognition motif for each target (86). Consequently, multiple applications have been found for targeted genome editing in experimental and preclinical models. Translating these applications to the clinic will, however, require thorough assessment of the off-target activity of the selected nuclease (87, 88) and optimization of the therapy. Gene addition by HDR or conventional vectors can also be combined with the targeted disruption of another cellular gene to augment effector function, such as by relieving inhibitory checkpoints when editing T-cell specificity (89).

Gene-editing strategies are also being developed for use in direct *in vivo* applications, including liver-directed gene therapy. Studies in mice have reported the reconstitution of FIX expression in haemophilia B mice by AAV-mediated delivery of ZFNs and a repair template (90), the targeted disruption of the cholesterol regulatory gene *Pcsk9* by AAV-mediated delivery of CRISPR/Cas9 (91), and the repair of the *Fah* mutation in tyrosinaemia (an inability to break down the amino acid tyrosine) by hydrodynamic plasmid delivery of CRISPR/Cas9 and a template (92). Major hurdles to the further development of *in vivo* gene editing include the safe and clinically suitable delivery of the editing machinery, which should act transiently without inducing cellular toxicity and immunogenicity. All artificial nucleases in current use employ one or more domains derived from prokaryotes - in some cases, common bacterial pathogens or saprophytes. Therefore, when delivered *in vivo*, there is a risk that such nucleases will generate or encounter a preexisting cell-mediated immunity, especially if they are expressed over the long term. More recently, an endonuclease-independent gene-targeting strategy was demonstrated in mice. A transgene was targeted to the albumin locus of hepatocytes by administering a promoterless AAV vector with homology to the highly transcribed albumin gene (93). A small proportion of the vector integrated at the albumin locus and became expressed from its endogenous promoter, which removed the requirement for both endonucleases and the transfer of a promoter within the vector. The efficiency and underlying mechanism of this and the other types of *in vivo* gene editing discussed in this Review are still to be fully determined.

Retinal gene therapy

The remarkable benefits that gene replacement can provide to patients with severe degenerative diseases were first highlighted by retinal gene therapy. Subretinal administration of AAV serotype 2 (AAV2)-mediated gene therapy in patients with type 2 Leber congenital amaurosis (LCA), an inherited retinal dystrophy that causes loss of vision at an early age, led to improved visual acuity in several

young patients in three independent trials (94-96). In two of these trials, the patients lost the benefit after a follow-up period of 2-3 years. However, sustained benefit was reported after a similar follow-up period in the third trial (97), which has now progressed to phase III testing. The reasons for the different outcome of these trials, all of which used an AAV2-derived capsid, remain unclear (98). The progressive degenerative nature of type 2 LCA poses a challenge for delivering extended therapeutic benefits because the small number of photoreceptors that are rescued by the therapy can eventually be overcome by non-cell-autonomous changes in the tissue. As discussed for liver-directed therapy, there might also be subtle differences in vector design and manufacturing that affect the extent of *in vivo* gene transfer, the inflammatory response at the site of delivery and the level of transgene reconstitution in transduced cells. The availability of AAV vectors with higher potencies, which would allow safer dose escalation and enhanced transduction, and more stringent tropism for the relevant targets, should help to overcome these limitations.

Other relevant developments

There are several other advances in the gene-therapy field that could not be discussed in detail in this Review. These include applications in neurodegenerative diseases that have reached the clinical-testing stage. For example, good safety but limited efficacy has been demonstrated for the delivery of transgenes to the brain by AAV or lentiviral vectors (99-101). Increasing the administered vector dose and optimizing the vehicle, cargo and study design are likely to lead to further advances.

Oncolytic viruses, which infect and kill cancer cells, have been in clinical use for some time. Although they can deliver robust and clinically relevant anticancer activity, these viruses are still being used as part of combination therapies (102). Oncolytic therapies exploit viral replication and the induction of an immune response against infected cells - conditions that are normally offset in gene-therapy strategies. Intriguingly, their efficacy is likely to be augmented by adding a transgene cargo that improves the induction of anticancer immunity (103).

Adenoviral vectors of simian origin are also being assessed for their ability to induce humoral and cellular immunity through vaccination: encouraging results have already been seen in emerging or widespread infectious diseases that have long resisted conventional attempts (104, 105). In preclinical models, the gene-based delivery of antibody therapy or prophylaxis is being explored to establish an *in vivo* stable and robust source of large quantity of antibodies with optimal specificity to treat or prevent infection (106).

Future outlook

Gene therapy could be poised to become an important new approach for the third millennium because its reach extends well beyond that of conventional drugs. Gene therapy enables the targeted delivery of information-rich gene-based cassettes that facilitate the stable, sustained and regulated expression of biological

agents. Furthermore, when combined with cell therapy, it turns cells into smart vehicles for targeted gene delivery. As exemplified in the studies discussed in this Review, gene therapy directs powerful biological processes towards the goals of disease correction, tissue repair and regeneration. For instance, the stability, fidelity and amplification of the delivered therapeutic can be guaranteed by transferring information by genetic mechanisms. The homing and trafficking mechanisms of cells in the human body can be used to target gene-based therapeutics to specific tissues and disease sites. Gene therapy also makes use of the regenerative potential of stem cells and transplantation as well as the biological weapon of immunity, which is exploited for the specific elimination of transformed or infected cells. By taking advantage of these inbuilt biological capabilities, gene therapy has the potential to address the substantial unmet medical needs of both rare and common severe diseases, which will benefit both patients and - more broadly - society.

Major challenges must still be addressed before this promise can be realized. For example, the efficacy and safety of gene-transfer vectors should be improved by further engineering their design and composition, which could include combining the biological features of different viruses with synthetic molecules. These advances will enable vectors to target tissues and cell types precisely (64, 63, 107), and overcome cellular restrictions on gene transduction and bypass sensors of exogenous nucleic acids. They will also help vectors to avoid activating the innate and adaptive immune system. Overall, the changes will also ensure that transgene expression is reproducible, robust, occurs over an extended period and closely mimics the endogenous pattern of expression (when gene replacement is performed). Improvements in vector manufacturing and characterization will allow the standardization and comparative assessment of vector performance between trials. The rate and specificity of *in situ* gene correction and editing, the integration of vectors at safe genomic harbours, and allele-specific silencing by artificial nucleases and epigenetic modifiers represent further opportunities for improving gene-therapy strategies. Deeper understanding of disease pathogenesis in inherited, multigenic or acquired conditions will enable the development of new gene-based treatment strategies. Because gene transfers employ 'live' biological drugs of unprecedented complexity that have the potential to induce extended effects on patients and their germ lines, long-term surveillance and precautionary measures must be taken while their use is being pioneered. Moreover, current - limited - understanding of the regulation of stem cells, tissue regeneration and immune-response checkpoints constrains the capacity for intervention and raises concerns about the untoward effects of manipulation.

From a clinical standpoint, the bedside delivery of gene and cell therapies calls for multidisciplinary expertise and, in some cases, advanced cell processing at the clinical-treatment site. Biological readouts must also be developed to monitor the safety and efficacy of therapies. As the first gene therapies progress from registration to marketing, both the pharmaceutical sector and regulatory agencies are being engaged to help define appropriate quality standards for manufacturing and release and to build suitable pipelines for supplying such highly personalized therapies. From a societal standpoint, the complexity and cost of manufacturing

and supplying 'live' biological drugs in conventional health-care systems will challenge the sustainability of these therapies and require creative cost-reimbursement policies that enable all patients to benefit from them (Box 3).

Finally, from an ethics standpoint, it is important to consider whether medicine should surrender to the rule of technology or commit to a more responsible steering of the course of progress. For instance, the avenues that are being opened to intervention by emerging technologies could undermine our self-perception and self-determination as we end up viewing ourselves as the evolutionary product of DNA that has become self-conscious and can edit itself to shape its progeny as and when it desires. The call for a moratorium on applying genome editing to human germline cells highlights forthcoming ethical dilemmas (108, 109). Science might turn again to the ancient roots of Western culture to learn from its wisdom, such as the "Know thyself" inscription on the Temple of Apollo at Delphi in ancient Greece a warning to recognize our limits. Modern philosophy has trained us to exercise criticism when assessing the truth and certainty of knowledge, which is limited a priori and enables learning relationships and making predictions but not uncovering the nature of things. Those predictions allow us to develop strategies that can alleviate human suffering from disease, thereby providing our endeavours with a translational framework that can guide our choices and justify them from an ethics standpoint.

BOX 3 - Pricing gene therapy

The first gene therapy to be commercially approved in the Western world was alipogene tiparvovec (also known as Glybera). This muscle-directed adeno-associated virus 1-based gene therapy was granted marketing authorization in the European Union in 2012 for the treatment of a rare form of familial dyslipidaemia (121). The market price was recently set at €1 million (US\$1.1 million) per treatment (122). If forthcoming gene therapies are also sold at such a high price, they will challenge the standard reimbursement policies of governments and insurance companies. This high price reflects the cost of preclinical development, manufacturing and distribution of the new medicine, especially for *ex vivo* gene therapies, which are highly personalized and require individualized manufacturing. However, gene therapies have the potential to deliver a substantial, long-lasting benefit to the patient on a single administration, which may offset the cost of the standard treatment of the condition and its complications. Nonetheless, a single upfront payment model for gene therapy may not be sustainable. Approaches that spread the payment over several years should be considered, which could be linked to the successful outcome of the therapy.

For example, a pay-for-performance strategy has been proposed (123) that makes the payment of instalments dependent on improved patient health, as determined by objective biomarkers. The risks are then shared between the health-care provider and insurer and the cost of treatment is more closely commensurate to the actual benefits delivered to the patient. In addition, the sustainability of gene therapies could be improved by adapting regulatory and manufacturing requirements to accommodate the unique features of these medicines and by facilitating their accessibility and distribution without detriment to their safety. A flexible platform-based approval and registration strategy should be considered, especially when developing gene and cell therapies that must be adapted to each individual, as is the case for the transfer of T-cell antigen receptor genes that target tumour neoantigens into autologous T cells of people with cancer (44-47).

Acknowledgements

L.N. apologizes to the many scientists whose contributions to the field could not be acknowledged owing to space limitations. He thanks past and present members of his laboratory and colleagues at the San Raffaele Telethon Institute for Gene Therapy (TIGET) and the San Raffaele Scientific Institute. L.N. is also grateful to the Telethon Foundation, the European Union (FP7 and ERC), the Italian Association for Cancer Research, and the Italian ministries of health and of scientific research for supporting his research.

Author Information

Reprints and permissions information is available at www.nature.com/reprints. The author declares competing financial interests: see go.nature/pnd2xw. Readers are welcome to comment on the online version of this paper at go.nature/pnd2xw. Correspondence should be addressed to L.N. (naldini.luigi@hsr.it).

Bibliografia

1. Naldini L. Ex vivo gene transfer and correction for cell-based therapies. *Nature Rev. Genet.* 2011; 12: 301-315.
2. Hacein-Bey-Abina S, et al. Efficacy of gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* 2010; 363: 355-364.
3. Aiuti A, et al. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *N. Engl. J. Med.* 2009; 360: 447-458.
4. Ferrua F, Brigida I, Aiuti A. Update on gene therapy for adenosine deaminase-deficient severe combined immunodeficiency. *Curr. Opin. Allergy Clin. Immunol.* 2010; 10: 551-556.
5. Fischer A, Hacein-Bey-Abina S, Cavazzana-Calvo M. 20 years of gene therapy for SCID. *Nature Immunol.* 11, 457-460 (2010). A comprehensive review of the therapeutic potential, risks and limitations of HSC-based SCID gene therapy using γ -RV by some of its pioneers; see also refs 3 and 7.
6. Gaspar HB, et al. Hematopoietic stem cell gene therapy for adenosine deaminase-deficient severe combined immunodeficiency leads to long-term immunological recovery and metabolic correction. *Sci. Transl. Med.* 2011; 3: 97ra80; erratum 2013; 5: 168er1.
7. Gaspar HB, et al. Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. *Sci. Transl. Med.* 2011; 3: 97ra79.
8. Boztug K, et al. Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *N. Engl. J. Med.* 2010; 363: 1918-1927.
9. Candotti F, et al. Gene therapy for adenosine deaminase-deficient severe combined immune deficiency: clinical comparison of retroviral vectors and treatment plans. *Blood.* 2012; 120: 3635-3646.
10. Kang EM, et al. Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils. *Blood.* 2010; 115: 783-791.

11. Hacein-Bey-Abina S, et al. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* 2008; 118: 3132-3142.
12. Howe, S. J. et al. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.* 2008; 118: 3143-3150.
13. Stein S, et al. Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nature Med.* 2010; 16: 198-204.
14. Braun CJ, et al. Gene therapy for Wiskott-Aldrich syndrome-long-term efficacy and genotoxicity. *Sci. Transl. Med.* 2014; 6: 227ra233.
15. Kang HJ, et al. Retroviral gene therapy for X-linked chronic granulomatous disease: results from phase I/II trial. *Mol. Ther.* 2011; 19: 2092-2101.
16. Aiuti A, et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science.* 2013; 341: 1233151. In this study, vector insertional analyses in patients show data consistent with improved safety of lentiviral vectors versus γ -RVs while achieving similarly effective disease correction; see also ref. 17.
17. Hacein-Bey Abina S, et al. Outcomes following gene therapy in patients with severe Wiskott-Aldrich syndrome. *J. Am. Med. Assoc.* 2015; 313: 1550-1563.
18. Hacein-Bey-Abina S, et al. A modified γ -retrovirus vector for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* 2014; 371: 1407-1417.
19. Cavazzana-Calvo M, et al. Transfusion independence and HMGA2 activation after gene therapy of human β -thalassaemia. *Nature.* 2010; 467: 318-322.
20. Cartier N, et al. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science.* 2009; 326: 818-823. The first trial of HSC gene therapy performed with lentiviral vectors shows data consistent with stable HSC transduction, with long-term safety and efficacy revealed in the follow-up paper (see ref. 21).
21. Cartier N, et al. Lentiviral hematopoietic cell gene therapy for X-linked adrenoleukodystrophy. *Methods Enzymol.* 2012; 507: 187-198.
22. Biffi A, et al. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science.* 2013; 341: 1233158. This study highlights the potential of genetic engineering by achieving the stable reconstitution of haematopoiesis in which up to 90% of cells are gene corrected and overexpress the transgene, which provides therapeutic benefit when conventional HSC transplantation is less satisfactory.
23. Notarangelo LD, et al. Primary immunodeficiencies: 2009 update. *J. Allergy Clin. Immunol.* 2009; 124: 1161-1178.
24. Kemp S, Berger J, Aubourg P. X-linked adrenoleukodystrophy: clinical, metabolic, genetic and pathophysiological aspects. *Biochim. Biophys. Acta.* 2012; 1822: 1465-1474.
25. Gieselmann V, Krageloh-Mann I. Metachromatic leukodystrophy—an update. *Neuropediatrics.* 2010; 41: 1-6.
26. Gennery AR, et al. Transplantation of hematopoietic stem cells and long-term

- survival for primary immunodeficiencies in Europe: entering a new century, do we do better? *J. Allergy Clin. Immunol.* 2010; 126: 602-610.
27. Krägeloh-Mann I, et al. Juvenile metachromatic leukodystrophy 10 years post transplant compared with a non-transplanted cohort. *Bone Marrow Transplant.* 2013; 48: 369-375.
 28. Copelan EA. Hematopoietic stem-cell transplantation. *N. Engl. J. Med.* 2006; 354: 1813-1826.
 29. Biffi A, et al. Correction of metachromatic leukodystrophy in the mouse model by transplantation of genetically modified hematopoietic stem cells. *J. Clin. Invest.* 2004; 113: 1118-1129.
 30. Biffi A, et al. Gene therapy of metachromatic leukodystrophy reverses neurological damage and deficits in mice. *J. Clin. Invest.* 2006; 116: 3070-3082.
 31. Capotondo A, et al. Brain conditioning is instrumental for successful microglia reconstitution following hematopoietic stem cell transplantation. *Proc. Natl Acad. Sci. USA.* 2012; 109: 15018-15023.
 32. Mingozzi F, High KA. Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nature Rev. Genet.* 2011; 12: 341-355.
 33. Nayak S, Herzo RW. Progress and prospects: immune responses to viral vectors. *Gene Ther.* 2010; 17: 295-304.
 34. Mingozzi F, High KA. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood.* 2013; 122: 23-36.
 35. Grieger JC, Samulski RJ. Adeno-associated virus vectorology, manufacturing, and clinical applications. *Methods Enzymol.* 2012; 507: 229-254.
 36. High KH, Nathwani A, Spencer T, Lillicrap D. Current status of haemophilia gene therapy. *Haemophilia.* 2014; 20 (Suppl. 4): 43-49.
 37. Berntorp E, Shapiro AD. Modern haemophilia care. *Lancet.* 2012; 379: 1447-1456.
 38. Manno CS, et al. Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nature Med.* 2006; 12: 342-347; erratum 2006; 12: 592. The first clinical data to show the safety and potential efficacy of liver-directed AAV gene transfer, which was unexpectedly abrogated by an immune response against viral capsids (as detailed in ref. 39).
 39. Mingozzi F, et al. CD8+ T-cell responses to adeno-associated virus capsid in humans. *Nature Med.* 2007; 13: 419-422.
 40. Nathwani AC, et al. Long-term safety and efficacy of factor IX gene therapy in hemophilia B. *N. Engl. J. Med.* 2014; 371: 1994-2004. This AAV8-based trial was first to report stable FIX expression at therapeutic levels and also first to overcome the detrimental effect of the immune response to viral capsids by corticosteroid administration.
 41. Rosenberg SA, Restif NP. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science.* 2015; 348: 62-68. A comprehensive review of the clinical development and potentially transformative impact of adoptive T-cell therapy on cancer by one of its pioneers; see also ref. 42.
 42. Maus MV, et al. Adoptive immunotherapy for cancer or viruses. *Annu. Rev. Immunol.* 2014; 32: 189-225.

43. Dudley ME, et al. Cancer regression and autoimmunity in patients after clonal repopulation with antitumor lymphocytes. *Science*. 2002; 298: 850-854.
44. Gubin MM, et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature*. 2014; 515: 577-581.
45. Yadav M, et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature*. 2014; 515, 572-576.
46. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science*. 2015; 348: 69-74. A timely review on the origin and nature of tumour neoantigens and how they can be identified and potentially exploited for targeted T-cell gene therapy in the clinical setting.
47. Rizvi NA, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015; 348: 124-128.
48. Hunder NN, et al. Treatment of metastatic melanoma with autologous CD4+ T cells against NY-ESO-1. *N. Engl. J. Med.* 2008; 358: 2698-2703.
49. Johnson LA, et al. Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen. *Blood*. 2009; 114: 535-546.
50. Robbins PF, et al. A pilot trial using lymphocytes genetically engineered with an NY-ESO-1-reactive T-cell receptor: long-term follow-up and correlates with response. *Clin. Cancer Res.* 2015; 21: 1019-1027.
51. Kochenderfer JN, et al. Chemotherapy-refractory diffuse large B-cell lymphoma and indolent B-cell malignancies can be effectively treated with autologous T cells expressing an anti-CD19 chimeric antigen receptor. *J. Clin. Oncol.* 2015; 33: 540-549.
52. Brentjens RJ, et al. CD19-targeted T cells rapidly induce molecular remissions in adults with chemotherapy-refractory acute lymphoblastic leukemia. *Sci. Transl. Med.* 2013; 5: 177ra38.
53. Maude SL, et al. Chimeric antigen receptor T cells for sustained remissions in leukemia. *N. Engl. J. Med.* 2014; 371: 1507-1517.
54. Lee DW, et al. T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial. *Lancet*. 2015; 385: 517-528.
55. June CH, Riddell SR, Schumacher TN. Adoptive cellular therapy: a race to the finish line. *Sci. Transl. Med.* 2015; 7: 280ps7.
56. Biffi A, et al. Lentiviral vector common integration sites in preclinical models and a clinical trial reflect a benign integration bias and not oncogenic selection. *Blood*. 2011; 117: 5332-5339.
57. Deichmann A, et al. Insertion sites in engrafted cells cluster within a limited repertoire of genomic areas after gammaretroviral vector gene therapy. *Mol. Ther.* 2011; 19: 2031-2039.
58. Doulatov S, Notta F, Laurenti E, Dick JE. Hematopoiesis: a human perspective. *Cell Stem Cell*. 2012; 10: 120-136.
59. Gattinoni L. Memory T cells officially join the stem cell club. *Immunity*. 2014; 41: 7-9.

60. Cieri N, et al. IL-7 and IL-15 instruct the generation of human memory stem T cells from naive precursors. *Blood*. 2013; 121: 573-584.
61. Biasco, L. et al. In vivo tracking of T cells in humans unveils decade-long survival and activity of genetically modified T memory stem cells. *Sci. Transl. Med.* 2015; 7: 273ra13.
62. Asokan A, Schaffer DV, Samulski RJ. The AAV vector toolkit: poised at the clinical crossroads. *Mol. Ther.* 2012; 20: 699-708.
63. Mingozzi F, et al. Overcoming preexisting humoral immunity to AAV using capsid decoys. *Sci. Transl. Med.* 2013; 5: 194ra92.
64. Lisowski L, et al. Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature*. 2014; 506: 382-386.
65. Kohn DB. Gene therapy outpaces haplo for SCID-X1. *Blood*. 2015; 125: 3521-3522.
66. Logan AC, Weissman IL, Shizuru JA. The road to purified hematopoietic stem cell transplants is paved with antibodies. *Curr. Opin. Immunol.* 2012; 24: 640-648.
67. Provasi E, et al. Editing T cell specificity towards leukemia by zinc finger nucleases and lentiviral gene transfer. *Nature Med.* 2012; 18: 807-815.
68. Torikai H, et al. A foundation for universal T-cell based immunotherapy: T cells engineered to express a CD19-specific chimeric-antigen-receptor and eliminate expression of endogenous TCR. *Blood*. 2012; 119: 5697-5705.
69. Sharma P, Allison P. The future of immune checkpoint therapy. *Science*. 2015; 348: 56-61.
70. Li H, et al. Assessing the potential for AAV vector genotoxicity in a murine model. *Blood*. 2011; 117: 3311-3319.
71. Chandler RJ, et al. Vector design influences hepatic genotoxicity after adeno-associated virus gene therapy. *J. Clin. Invest.* 2015; 125: 870-880.
72. Nault J-C, et al. Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. *Nature Genet.* 2015; <http://dx.doi.org/10.1038/ng.3389>.
73. Martino AT, et al. Engineered AAV vector minimizes in vivo targeting of transduced hepatocytes by capsid-specific CD8+ T cells. *Blood*. 2013; 121: 2224-2233.
74. Kotterman MA, Schaffer DV. Engineering adeno-associated viruses for clinical gene therapy. *Nature Rev. Genet.* 2014; 15: 445-451.
75. Cantore A, et al. Liver-directed lentiviral gene therapy in a dog model of hemophilia B. *Sci. Transl. Med.* 2015; 7: 277ra28.
76. Cox DB, Platt RJ, Zhang F. Therapeutic genome editing: prospects and challenges. *Nature Med.* 2015; 21: 121-131.
77. Tebas P, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* 2014; 370: 901-910. The first clinical testing of targeted gene disruption that showed the safety, persistence and survival advantage of T cells that have been genetically edited for resistance to HIV-1.
78. Li L, et al. Genomic editing of the HIV-1 coreceptor CCR5 in adult hemato-

- poietic stem and progenitor cells using zinc finger nucleases. *Mol. Ther.* 2013; 21: 1259-1269.
79. Bauer DE, et al. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science.* 2013; 342: 253-257.
80. Lombardo A, et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nature Methods.* 2011; 8: 861-869.
81. Rio P, et al. Targeted gene therapy and cell reprogramming in Fanconi anemia. *EMBO Mol. Med.* 2014; 6: 835-848.
82. Genovese P, et al. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature.* 2014; 510: 235-240. This paper demonstrates differential permissiveness to targeted genome editing in haematopoietic stem and progenitor cells and provides a proof of concept for the in situ correction of SCID-X1 mutations in HSCs.
83. Hoban MD, et al. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood.* 2015; 125: 2597-2604.
84. Osborn MJ, et al. Fanconi anemia gene editing by the CRISPR/Cas9 system. *Hum. Gene Ther.* 2015; 26: 114-126.
85. Jasin M, Rothstein R. Repair of strand breaks by homologous recombination. *Cold Spring Harb. Perspect. Biol.* 2013; 5: a012740.
86. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science.* 2014; 346: 1258096. In this review, the researchers who pioneered the application of RNA-guided nucleases to genome engineering show how this transformative technique can make targeted genome editing easy.
87. Gabriel R, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nature Biotechnol.* 2011; 29: 816-823.
88. Tsa SQ, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nature Biotechnol.* 2015; 33: 187-197.
89. Beane JD, et al. Clinical scale zinc finger nuclease-mediated gene editing of PD-1 in tumor infiltrating lymphocytes for the treatment of metastatic melanoma. *Mol. Ther.* 2015; 23: 1380-1390.
90. Li H, et al. In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature.* 2011; 475: 217-221. The first study to show the feasibility of targeted genome editing in vivo by the AAV-mediated delivery of artificial nucleases and template.
91. Ran FA, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature.* 2015; 520: 186-191.
92. Yin H, et al. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nature Biotechnol.* 2014; 32: 551-553.
93. Barzel A, et al. Promoterless gene targeting without nucleases ameliorates haemophilia B in mice. *Nature.* 2015; 517: 360-364.
94. Simonelli F, et al. Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration. *Mol. Ther.* 2010; 18: 643-650.
95. Jacobson SG, et al. Improvement and decline in vision with gene therapy in childhood blindness. *N. Engl. J. Med.* 2015; 372: 1920-1926.

96. Bainbridge JW, et al. Long-term effect of gene therapy on Leber's congenital amaurosis. *N. Engl. J. Med.* 2015; 372: 1887-1897.
97. Testa F, et al. Three-year follow-up after unilateral subretinal delivery of adeno-associated virus in patients with Leber congenital amaurosis type 2. *Ophthalmology.* 2013; 120: 1283-1291.
98. Wright AF. Long-term effects of retinal gene therapy in childhood blindness. *N. Engl. J. Med.* 2015; 372: 1954-1955.
99. Leone P, et al. Long-term follow-up after gene therapy for canavan disease. *Sci. Transl. Med.* 2012; 4: 165ra163.
100. Tardieu M, et al. Intracerebral administration of adeno-associated viral vector serotype rh.10 carrying human SGSH and SUMF1 cDNAs in children with mucopolysaccharidosis type IIIA disease: results of a phase I/II trial. *Hum. Gene Ther.* 2014; 25: 506-516.
101. Palfi S, et al. Long-term safety and tolerability of ProSavin, a lentiviral vector-based gene therapy for Parkinson's disease: a dose escalation, open-label, phase 1/2 trial. *Lancet.* 2014; 383: 1138-1146.
102. Miest TS, Cattaneo R. New viruses for cancer therapy: meeting clinical needs. *Nature Rev. Microbiol.* 2014; 12: 23-34.
103. Lichty BD, Breitbach CJ, Stojdl DF, Bell JC. Going viral with cancer immunotherapy. *Nature Rev. Cancer.* 2014; 14: 559-567.
104. Ogwang C, et al. Prime-boost vaccination with chimpanzee adenovirus and modified vaccinia Ankara encoding TRAP provides partial protection against *Plasmodium falciparum* infection in Kenyan adults. *Sci. Transl. Med.* 2015; 7: 286re5.
105. Rampling T, et al. A monovalent chimpanzee adenovirus ebola vaccine - preliminary report. *N. Engl. J. Med.* 2015; <http://dx.doi.org/10.1056/NEJMoa1411627>.
106. Balaz AB, et al. Vectored immunoprophylaxis protects humanized mice from mucosal HIV transmission. *Nature Med.* 2014; 20: 296-300.
107. Girard-Gagnepain A, et al. Baboon envelope pseudotyped LVs outperform VSV-G-LVs for gene transfer into early-cytokine-stimulated and resting HSCs. *Blood.* 2014; 124: 1221-1231.
108. Baltimore D, et al. Biotechnology. A prudent path forward for genomic engineering and germline gene modification. *Science.* 2015; 348: 36-38.
109. Bosley KS, et al. CRISPR germline engineering - the community speaks. *Nature Biotechnol.* 2015; 33: 478-486.
110. Baxter. Baxalta reports continued progress on phase 1/2 clinical trial of BAX335, investigational gene therapy treatment for hemophilia B. *Baxter.* 2015; http://www.baxter.com/news-media/newsroom/press-releases/2015/06_24_15_bax335.page.
111. Montini E, et al. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nature Biotechnol.* 2006; 24: 687-696.
112. Modlich U, et al. Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood.* 2006; 108: 2545-2553.

113. Zychlinski D, et al. Physiological promoters reduce the genotoxic risk of integrating gene vectors. *Mol. Ther.* 2008; 16: 718-725.
114. Montini E, et al. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J. Clin. Invest.* 2009; 119: 964-975. This preclinical study highlights important features of vector design that affect genotoxicity and reveals strategies to alleviate it; the study was instrumental in promoting the clinical testing of improved vectors (see refs 112-113 for an in vitro assay that provides complementary information).
115. Modlich U, et al. Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol. Ther.* 2009; 17: 1919-1928.
116. Zhou S, et al. A self-inactivating lentiviral vector for SCID-X1 gene therapy that does not activate LMO2 expression in human T cells. *Blood.* 2010; 116: 900-908.
117. Zhou S, et al. Mouse transplant models for evaluating the oncogenic risk of a self-inactivating XSCID lentiviral vector. *PLoS One.* 2013; 8: e62333.
118. Baum C, Modlich U, Gohring G, Schlegelberger B. Concise review: managing genotoxicity in the therapeutic modification of stem cells. *Stem Cells.* 2011; 29: 1479-1484.
119. Amendola M, Venneri MA, Biffi A, Vigna E, Naldini L. Coordinate dual-gene transgenesis by lentiviral vectors carrying synthetic bidirectional promoters. *Nature Biotechnol.* 2005; 23: 108-116.
120. Greco R, et al. Improving the safety of cell therapy with the TK-suicide gene. *Front. Pharmacol.* 2015; 6: 95.
121. Melchiorri D, et al. Regulatory evaluation of Glybera in Europe - two committees, one mission. *Nature Rev. Drug Discov.* 2013; 12: 719.
122. Morrison C. \$1-million price tag set for Glybera gene therapy. *Nature Biotechnol.* 2015; 33: 217-218.
123. Brennan TA, Wilson JM. The special case of gene therapy pricing. *Nature Biotechnol.* 2014; 32: 874-876.

The epigenetic dimension of cancer: between development and plasticity

Pietro Lo Riso

Department of Experimental Oncology, Istituto Europeo di Oncologia, Milano

Cancer has been classically regarded as a multifaceted disease initiated by genetic mutations of oncogenes or tumor suppressors, that help transformed cells to escape proliferation restrictions and to achieve a growth advantage over the surrounding cellular environment. The selection of features that increase the fitness of cells result in the progressive acquisition of the so called “hallmarks of cancer”, that span several crucial aspects of cellular survival and result also in the hijacking of host’s immune and angiogenic system (1, 2). Thus, it is clear that several aspects must be taken into account in order to achieve the optimal therapeutic regime for patients, likely through a combinatorial approach that is able to target several hallmarks at once in order to avoid tumor escape.

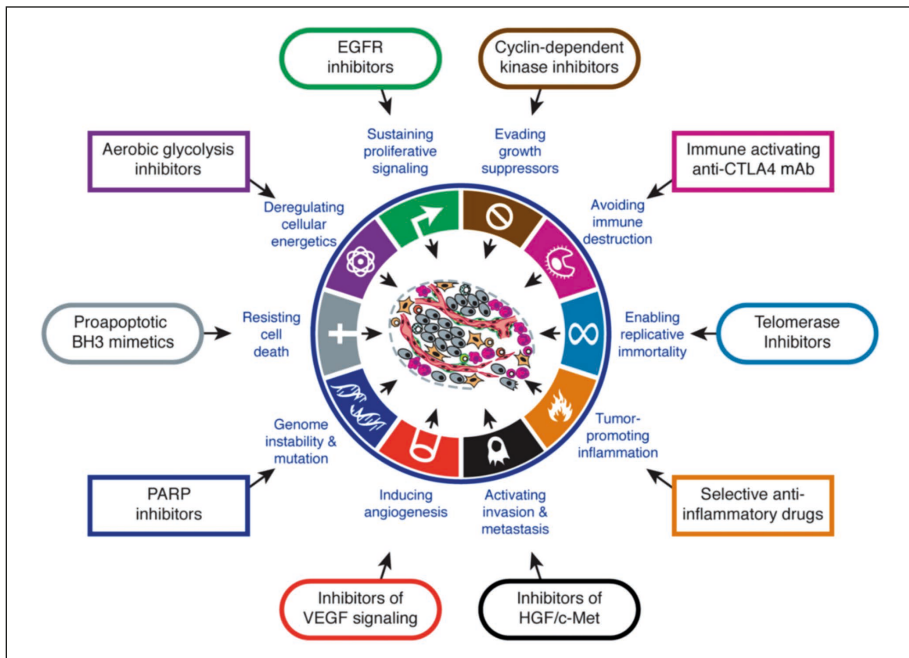


Fig. 1 - The hallmarks of cancer: a combinatorial approach (Hanahan and Weinberg 2011).

In the last years, extensive effort has been put to identify novel driver mutations in each primary tumor, underlying the diffused transcriptional dysregulation that characterize this disease (TCGA consortium 2011-2018, MSK-IMPACT 2017). In some cases, common mutations between histologically different tumors have been found, thus pointing to the possibility to use the same therapy to target the same mutations in different tumors, the so-called basket clinical trials. These trials however showed that targeting the same mutation in different tumors was not universally efficient, such as in the case of the treatment with Vemurafenib of V600 BRAF mutated tumors. This treatment, indeed, was efficient in a few tumor-types but resulted in anecdotal responses in the largest majority of tumors (3), thus pointing to a more complex interaction between single mutations with different mutated genomes and alluding to an influence of the transcriptional/epigenetic landscape of each cell type.

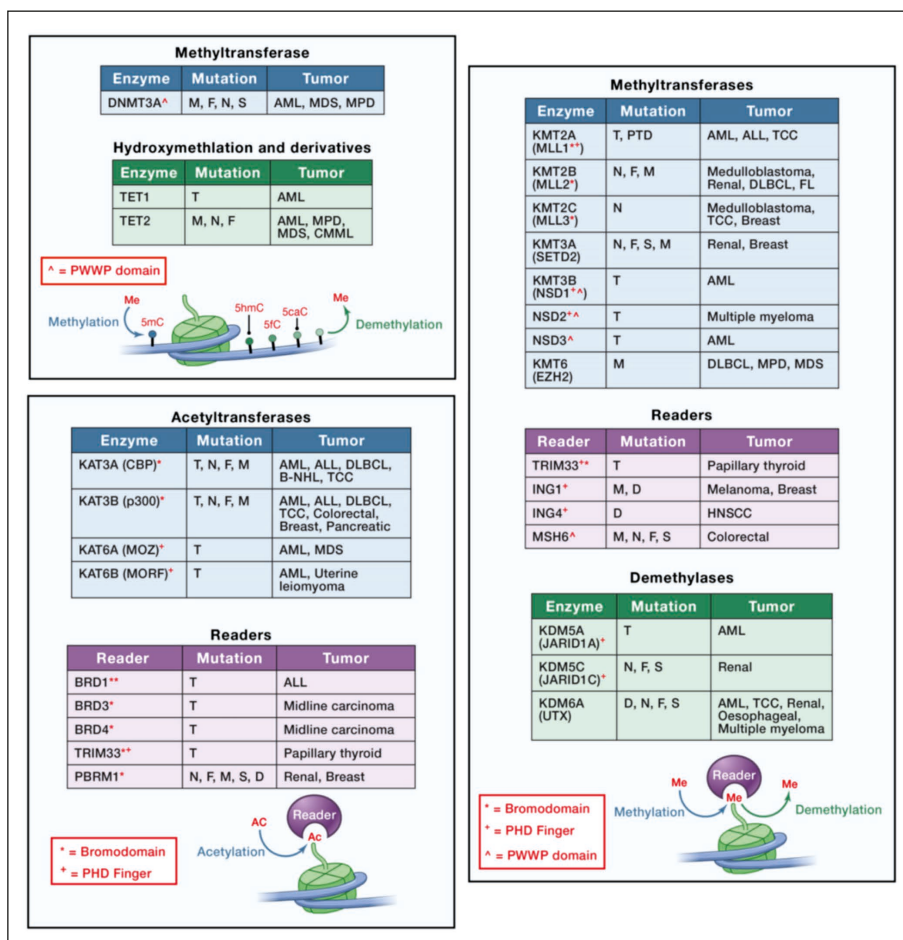


Fig. 2 - Epigenetic modifiers mutated in cancer (Dawson and Kouzarides 2012).

More recently, a number of evidence highlighted an important role of epigenetically-mediated transcriptional dysregulation in cancer progression (4). Indeed, several epigenetic marks that are known to have an impact on gene expression are often found deregulated in several cancer types (5-7). Also, several mutations have been found in epigenetic modifier genes (8).

Importantly, a new driving role for epigenetics has been recently found in pediatric ependymomas, where low to no DNA mutations were found, while a CpG methylator phenotype was present (9).

These findings put epigenetics in the spotlight as additional key players in tumorigenesis.

Given this important role in gene regulation and tumor progression, a number of drugs targeting epigenetic modifiers, the so-called epi-drugs, have been developed, showing very good results especially in the treatment of hematopoietic malignancies (10). Some of them have been also approved from FDA for clinical testing.

Epigenetic marks can be used also as trackers during tumor evolution and to define cell identity.

Specifically, DNA methylation is a very stable mark that has been used successfully both to define the parental tissue of origin in metastatic tumors of unknown origin (11) and to classify tumors of the central nervous system (12).

A further demonstration that the epigenetic *milieu* can affect the gene expression pattern is the demonstration that some genetic alterations can promote tumorigenesis only in the context of specific cell types. Indeed, not all the possible mutations are present in all kinds of tumors, thus pointing to a cell type-dependency of their tumorigenic potential.

The possibility to reprogram cells to a pluripotent state and to differentiate them to whatever cell type, paved the way to studies aimed at the dissection of the impact of cell type-specific epigenetic and transcriptional landscapes to genetic mutations.

One of the first demonstrations that mutations can unleash their tumorigenic effects only in specific contexts came with the reprogramming by somatic cell nuclear transfer of a RAS overexpressing melanoma genome (13). The derived pluripotent cell was indistinguishable from a normal one, thus showing that the pluripotent epigenetic landscape was able to suppress the malignant phenotype.

Drugs	classification	Approved Year	Indicated disease	ORR
Azacitidine	DNMT inhibitor	2004	MDS	17.9%
Vorinostat	HDAC inhibitor	2006	CTCL	30%
Decitabine	DNMT inhibitor	2006	MDS	42%-54%
Romidepesin	HDAC inhibitor	2009	TCL	34%
Ruxolitinib	JAK1/2 inhibitor	2011	Myelofibrosis	30%
Belinostat	HDAC inhibitor	2015	PTCL	25.8%
Panobinostat	HDAC inhibitor	2015	MM	NA

MDS, Myelodysplastic syndrome; CTCL, Cutaneous T-cell lymphoma; TCL, T-cell lymphoma; PTCL, Peripheral T-cell lymphoma; MM, Multiple Myeloma, ORR, Objective response rate.

Tab. 1 - FDA-approved epi-drugs (Li et al. 2017).

When used to generate viable mice, that carried in every cell in the body the mutated gene, these mice developed only melanomas and head and neck tumors. This set of experiments proved that the epigenetic landscape of cells can suppress the tumorigenic phenotype, that is evident only in a cell-type dependent way.

With the advent of induced pluripotent stem cells (iPSC) (14), this experimental could be scaled up *in vitro* and applied to many different tumors of human origin. The outcomes were dependent on the considered tumors. The main findings confirmed indeed that the pluripotent stage is in most cases tolerant to mutations and that the tumorigenic potential is dependent on the developmental lineage, but the result of differentiation could be:

- 1) tumors indistinguishable from the parental ones when differentiated in disease relevant lineages, such as in the case of glioblastoma (15) and chronic myeloid leukemia (16, 17);
- 2) tumors recapitulating the early steps of tumorigenesis then evolving to their most aggressive form, such as in the case of pancreatic ductal adenocarcinoma (18);
- 3) cells with very low tumorigenic potential (19-21).

These evidence show that the epigenetic resetting operated by reprogramming (22), is able to remove, at least in some cases, aberrant epigenetic marks that are fundamental for tumor establishment and/or progression. Thus, this approach can be used to investigate and dissect the epigenetic vs the genetic contribution to tumorigenesis.

Also, the reprogramming process and tumorigenesis share a number of interesting features. To this regard, it is known that poorly differentiated tumors show expression of pluripotency markers (23-25) and is usually associated with poor prognosis (26, 27).

Interestingly, the transient expression of the reprogramming factors in mice lead to metaplasia of renal tissues and, if more sustained, to the onset of Wilms tumors (28), thus pointing to a role of the reprogramming factors in cellular plasticity and tumorigenesis.

Case study: High Grade Serous Ovarian Cancer

High Grade Serous Ovarian Cancer (HGSOC) accounts for more than 70% of ovarian neoplasias. It is a major cause of cancer-related mortality, with really minor improvements on patient's survival achieved during the last 30 years (29, 30).

In vivo and *in vitro* models of HGSOC fail to recapitulate both advanced and early disease stages (30, 31), thus pointing to the acute need of new models for the accurate identification of disease-relevant molecular aberrations.

Moreover, there is still much controversy regarding the attribution to HGSOC of the correct cell of origin, either the fallopian tube epithelium (FTE) or the ovarian surface epithelium (OSE) (32). To address these issues, our lab has adopted two different approaches. The first one is based on the identification of an epigenetic blueprint that could be conserved in the transition between normal and tumor tissues, thus allowing to solve the attribution of the correct cell of origin. The second one aims at the establishment of a new *in vitro* model based on iPSC

to identify both early tumorigenic mechanisms and new targets for therapy from a developmental perspective.

The developmental origin of HGSOc

The main aim of this project was to define a molecular blueprint of the two tissues of origin that could be conserved upon tumor transformation, thus allowing a retrospective stratification of samples. We resorted to DNA methylation, which has been successfully used as a solid marker to be followed either during differentiation of various tissues or to stratify tumors based on their origin (11, 12, 33). Indeed, we could show that DNA methylation could be used also to track down the original compartment from which tumor arises.

In particular, we studied the global DNA methylation pattern of both FTE and OSE. We could show that a DNA methylation *blueprint* of these normal tissues is indeed preserved upon tumor transformation, thus allowing to bipartition the previously general category of HGSOc into FTE-like and OSE-like tumors.

Moreover, we could show that this distinction is the principal source of variance in DNA methylation between samples, thus pointing to the existence of a relevant biological difference between these two subgroups.

Then, we profiled the newly classified tumors by RNAseq and derived differentially expressed genes (DEGs) between FTE-like and OSE-like tumors. By these DEGs we could stratify a well-characterized retrospective cohort of HGSOc, highlighting a neat difference in survival between the two groups. Thus, the tissue of origin impacts the evolution of the tumor, with visible effects also on the fitness of patients.

This approach allowed to redefine the once homogenous pathological category of HGSOc into two clinically relevant subgroups, thus allowing to compare at the molecular level each tumor to the correct normal tissue of origin. This will allow to define druggable targets that are specific for each subset, thus improving patient's care in the context of precision medicine.

Reprogramming to pluripotency of HGSOc

Another approach we are currently using to highlight early molecular dysregulations in HGSOc is aimed at the resetting of the epigenetic landscape of these tumors. In this context, we are using reprogramming to pluripotency as a tool to:

- 1) derive iPSC from tumors and patient-matched skin fibroblasts, that carry the genomic landscape of the parental tumor with a reset epigenomic landscape;
- 2) verify whether the effect of the underlying mutations is already evident at any pluri- or multipotency stage;
- 3) differentiate these cells towards FTE and OSE to unleash the potential of genomic alterations in the epigenetic context that gave rise to the full-fledged tumor *in vivo*.

We could show that upon reprogramming and differentiation, neither at the pluripotent nor at the early mesodermal progenitor stage (the latter being the common developmental origin of both FTE and OSE) the impact of the mutated ge-

nome is translated into transcriptomic alterations, thus pointing to a pathogenetic mechanism for HGSOC that relies on context-specific epigenetic regulatory landscapes. It is thus likely that, by further differentiating these cells in the context of the two tissues of origin, it will be possible to shed light on the forerunning events of HGSOC pathogenesis.

These two complementary methodologies, thus, show in practice the potential of epigenetic studies in highlighting clinically relevant features of cancer.

Bibliografia

1. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000; 100: 57-70.
2. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144: 646-674.
3. Hyman DM, Puzanov I, Subbiah V, et al. Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. *N Engl J Med*. 2015; 373: 726-736.
4. Timp W, Feinberg AP. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. *Nat Rev Cancer*. 2013; 13: 497-510.
5. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet*. 2007; 8: 286-298.
6. Shen H, Laird PW. Interplay between the cancer genome and epigenome. *Cell*. 2013; 153: 38-55.
7. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell*. 2012; 22: 9-20.
8. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. *Cell*. 2012; 150: 12-27.
9. Mack SC, Witt H, Piro RM, et al. Epigenomic alterations define lethal CI-MP-positive ependymomas of infancy. *Nature*. 2014; 506: 445-450.
10. Li J, Hao D, Wang L, et al. Epigenetic targeting drugs potentiate chemotherapeutic effects in solid tumor therapy. *Sci Rep*. 2017; 7: 4035.
11. Moran S, Martinez-Cardus A, Sayols S, et al. Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol*. 2016; 17: 1386-1395.
12. Capper D, Jones DTW, Sill M, et al. DNA methylation-based classification of central nervous system tumours. *Nature*. 2018; 555: 469-474.
13. Hochedlinger K, Blalock R, Brennan C, et al. Reprogramming of a melanoma genome by nuclear transplantation. *Genes Dev*. 2004; 18: 1875-1885.
14. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006; 126: 663-676.
15. Stricker SH, Feber A, Engstrom PG, et al. Widespread resetting of DNA methylation in glioblastoma-initiating cells suppresses malignant cellular behavior in a lineage-dependent manner. *Genes Dev*. 2013; 27: 654-669.
16. Carette JE, Pruszk J, Varadarajan M, et al. Generation of iPSCs from cultured human malignant cells. *Blood*. 2010; 115 : 4039-4042.

17. Kumano K, Arai S, Kurokawa M. Generation of iPS cells from normal and malignant hematopoietic cells. *Int J Hematol.* 2013; 98: 145-152.
18. Kim J, Hoffman JP, Alpaugh RK, et al. An iPSC line from human pancreatic ductal adenocarcinoma undergoes early to invasive stages of pancreatic cancer progression. *Cell Rep.* 2013; 3: 2088-2099.
19. Mahalingam D, Kong CM, Lai J, Tay LL, Yang H, Wang X. Reversal of aberrant cancer methylome and transcriptome upon direct reprogramming of lung cancer cells. *Sci Rep.* 2012; 2: 592.
20. Miyoshi N, Ishii H, Nagai K, et al. Defined factors induce reprogramming of gastrointestinal cancer cells. *Proc Natl Acad Sci USA.* 2010; 107: 40-45.
21. Zhang X, Cruz FD, Terry M, Remotti F, Matushansky I. Terminal differentiation and loss of tumorigenicity of human cancers via pluripotency-based reprogramming. *Oncogene.* 2013; 32: 2249-2260, 2260 e2241-2221.
22. Papp B, Plath K. Reprogramming to pluripotency: stepwise resetting of the epigenetic landscape. *Cell Res.* 2011; 21: 486-501.
23. Ben-Porath I, Thomson MW, Carey VJ, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet.* 2008; 40: 499-507.
24. Chiou SH, Wang ML, Chou YT, et al. Coexpression of Oct4 and Nanog enhances malignancy in lung adenocarcinoma by inducing cancer stem cell-like properties and epithelial-mesenchymal transdifferentiation. *Cancer Res.* 2010; 70: 10433-10444.
25. Jeter CR, Liu B, Liu X, et al. NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene.* 2011; 30: 3833-3845.
26. Meng HM, Zheng P, Wang XY, et al. Over-expression of Nanog predicts tumor progression and poor prognosis in colorectal cancer. *Cancer Biol Ther.* 2010; 9: 295-302.
27. Nagata T, Shimada Y, Sekine S, et al. Prognostic significance of NANOG and KLF4 for breast cancer. *Breast Cancer.* 2014; 21: 96-101.
28. Ohnishi K, Semi K, Yamamoto T, et al. Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. *Cell.* 2014; 156: 663-677.
29. Bowtell DD. The genesis and evolution of high-grade serous ovarian cancer. *Nat Rev Cancer.* 2010; 10: 803-808.
30. Vaughan S, Coward JI, Bast RC, Jr., et al. Rethinking ovarian cancer: recommendations for improving outcomes. *Nat Rev Cancer.* 2011; 11: 719-725.
31. Bowtell DD, Bohm S, Ahmed AA, et al. Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer. *Nat Rev Cancer.* 2015; 15: 668-679.
32. Ng A, Barker N. Ovary and fimbrial stem cells: biology, niche and cancer origins. *Nat Rev Mol Cell Biol.* 2015; 16: 625-638.
33. Kim M, Costello J. DNA methylation: an epigenetic mark of cellular memory. *Exp Mol Med.* 2017; 49: e322.

Experimental epidemiology: neurodevelopmental epigenetics across human genetic variability

Nicolò Caporale

Department of Experimental Oncology, Istituto Europeo di Oncologia, Milano

The advent of induced pluripotent stem cells (iPSC) (1) allowed the scientific community to explore the molecular mechanisms underlying human brain pathophysiology in developmental stages and resolution that was not thinkable before.

iPSC are a self renewable system that can be derived from easy accessible human fibroblasts and indefinitely expanded *in vitro*. Moreover they can be differentiated into any cell lineage. As a matter of fact, a number of protocol flourished in the last decade, as an attempt to recapitulate the development of the human brain, allowing a more comprehensive understanding of the cellular and molecular landscape of the developing human brain (2).

Those technological advances are determining a great improvement of stem cell based disease modeling, for the opportunity to explore the molecular mechanisms that are dysregulated in both genetically and environmentally caused neurodevelopmental disorders (3).

In particular, we can investigate the molecular mechanisms that underlie the increased risk of developing intellectual disabilities caused by the interactions between early life exposure to endocrine disruptive chemical mixtures and different genetic backgrounds.

Human populations are exposed to a large number of chemicals with endocrine disrupting properties (EDCs) (4).

Their regulation represents a major unmet challenge due to the fact that while exposure to single EDCs has repeatedly been associated with major disorders and impaired development (5), real life entails simultaneous exposure to multiple EDCs in mixtures, with additive effects at lower doses than experimental effect thresholds for single compounds (6).

An integrated epidemiological-experimental design assessing the impact of EDC mixtures on human health and development

Experimental evidence with mixtures is most often limited to combinations within the same chemical class or to observational measurements on more complex mixtures, thus lacking causative weight to link actual population-based ex-

posure with adverse health outcomes in humans. We pursued a systematic integration of epidemiological and experimental evidence to elucidate the molecular pathways affected by EDC mixtures that are causally related to adverse outcomes in humans.

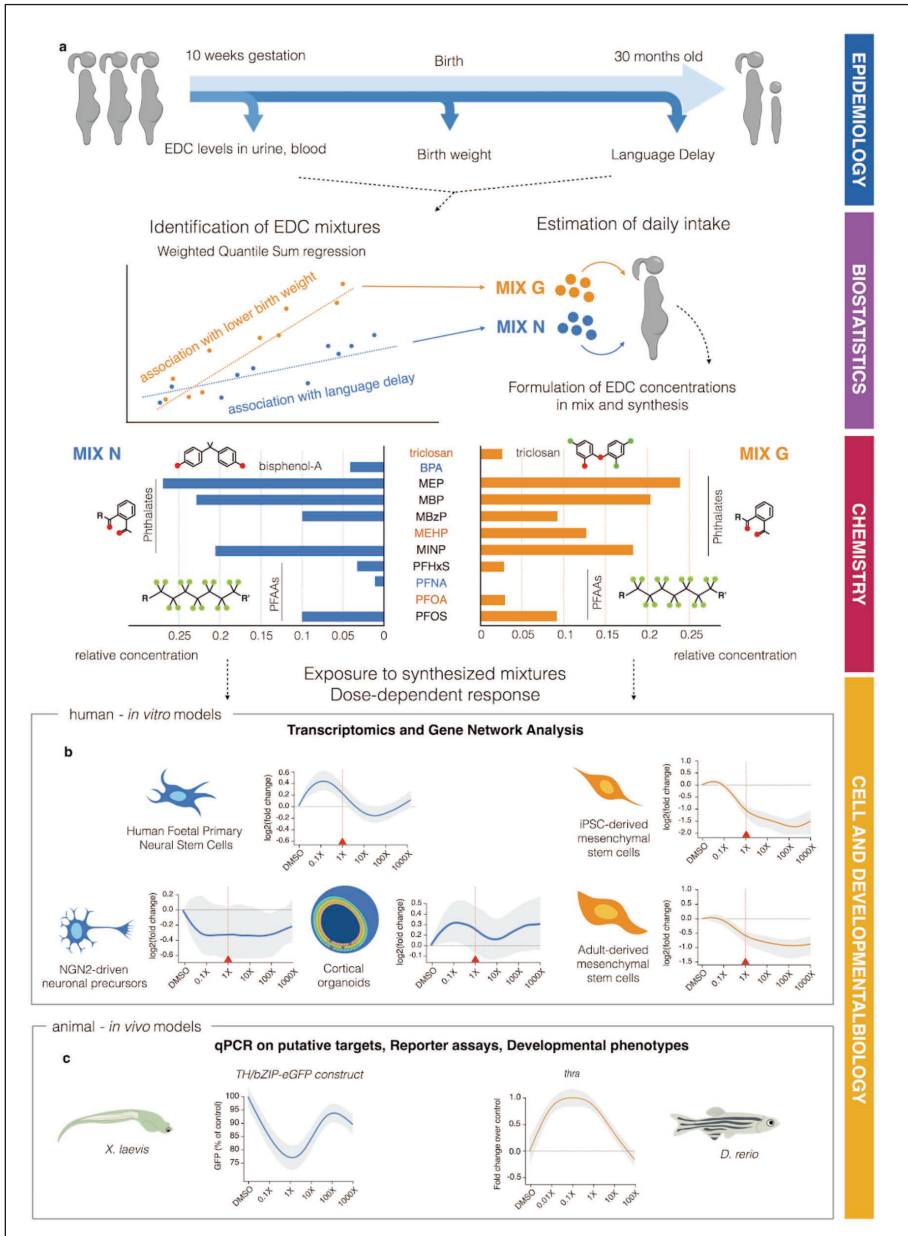


Fig. 1

To assess health outcomes of real-life EDC exposures we harnessed:

- 1) the power of a population-based mother-child pregnancy cohort to measure prenatal EDC exposures, combined with novel biostatistical tools to infer associations between specific EDC mixtures and two child health domains, neurodevelopment and metabolism/growth (Figure 1a);
- 2) complementary assays in human systems, to establish causality and deconvolute gene regulatory networks and cellular responses dysregulated by these EDC mixtures in concentrations corresponding to human exposure (Figure 1b);
- 3) paradigmatic *in vivo* models to determine the physiological impact of key affected pathways.

MIX N disrupts human neurodevelopmental pathways

To define the molecular impact of MIX N, we employed primary neural stem cells sourced from cortex and ganglionic eminence of human foetuses (Figure 2a) at post-conception week (PCW) 11 and 8, respectively (henceforth Human Foetal Primary Neural Stem Cells (HFPNSC)). Given the potentially non-linear and non-monotonic dose-response patterns associated with EDC mixtures (7, 7), the experimental design included 5 doses of MIX N, ranging from 0.1X to 1000X and a global assessment of impact on gene expression. To this end, RNA-seq was performed after 48 h MIX N exposure and patterns of EDC dose-dependent transcriptional responses determined using an analysis that considers MIX N dilutions (including the DMSO control) as distinct categories. This unbiased approach, which does not assume any particular response pattern (e.g. linearity or monotony), allowed us to define lists of differentially-expressed genes (DEGs, henceforth 'unbiased DEGs'), which were subsequently clustered on the basis of their dose-response patterns (Figure 2b). The dose-response patterns showed dysregulation already at low concentrations, highlighting the significance of doses recapitulating human exposure.

Functional characterization of DEGs revealed enrichment in Gene Ontology (GO) categories related to chromatin modulation and regulation of gene expression (Figure 2c), showing a major and specific impact of EDC during early fore-brain development. Given the epidemiological evidence linking MIX N exposure to verbal competence, and the central role of chromatin dysfunction in autism and intellectual disabilities (9), we tested whether EDC-induced DEGs were enriched for genes associated with these conditions. We found significant enrichment for genes associated with Intellectual Disabilities ($p \approx 4.7 \times 10^{-14}$ ID (10)), Developmental Disorders ($p \approx 1.5 \times 10^{-6}$ DDD (11)) and Autism Spectrum Disorders (ASD) ($p \approx 5.1 \times 10^{-5}$ in the Autism Speaks-Google MSSNG database (12) and $p \approx 2.7 \times 10^{-4}$ in the Autism Spectrum/Intellectual Disability (ASID) database (13)) (Figure 2c).

While primary neural stem cells directly sourced from fetal telencephali represent the arguably most proximal model of human neurodevelopment, their availability is limited and hence they are ill suited for large-scale and iterative studies required to advance regulatory toxicology. We thus validated our transcriptomic findings on MIX N using neurodevelopmental models based on self-renewing

sources of human induced pluripotent stem cells (iPSC): 3D cortical organoids that recapitulate human *in vivo* corticogenesis (14).

Immunofluorescence confirmed expression of specific markers, including PAX6 and Nestin as well as the defining arrangement in ventricular zone-like structures lined by ZO-1 expressing cells (Figure 2d).

Importantly, organoids largely recapitulated the dose-response patterns seen for DEGs in HFPNSC (compare Figure 2e with 2b), thus representing a potentially transforming tool for large-scale regulatory toxicology.

Our validation of human iPSC-based neurodevelopmental models for assessing EDCs paves the way to the systematic evaluation of EDC toxicity across different human genetic backgrounds using representative iPSC collections from defined populations, for which we empirically derived the optimal transcriptome-based disease modeling designs (15). Since MIX N was defined by verbal proficiency we thus reasoned that a genetic background conferring relative resilience of verbal skills would provide a rigorous proof of principle of the applicability of our findings across genetic backgrounds. To this end, we derived cortical organoids from iPSC lines harboring the 7q11.23 hemi-deletion that causes Williams-Beuren syndrome (WBS), a neurodevelopmental disorder characterized by cognitive weaknesses that selectively spare language abilities, and for which we previously uncovered disease-relevant dysregulation in iPSC and neural progenitors (16). Using the same analytical approach outlined above

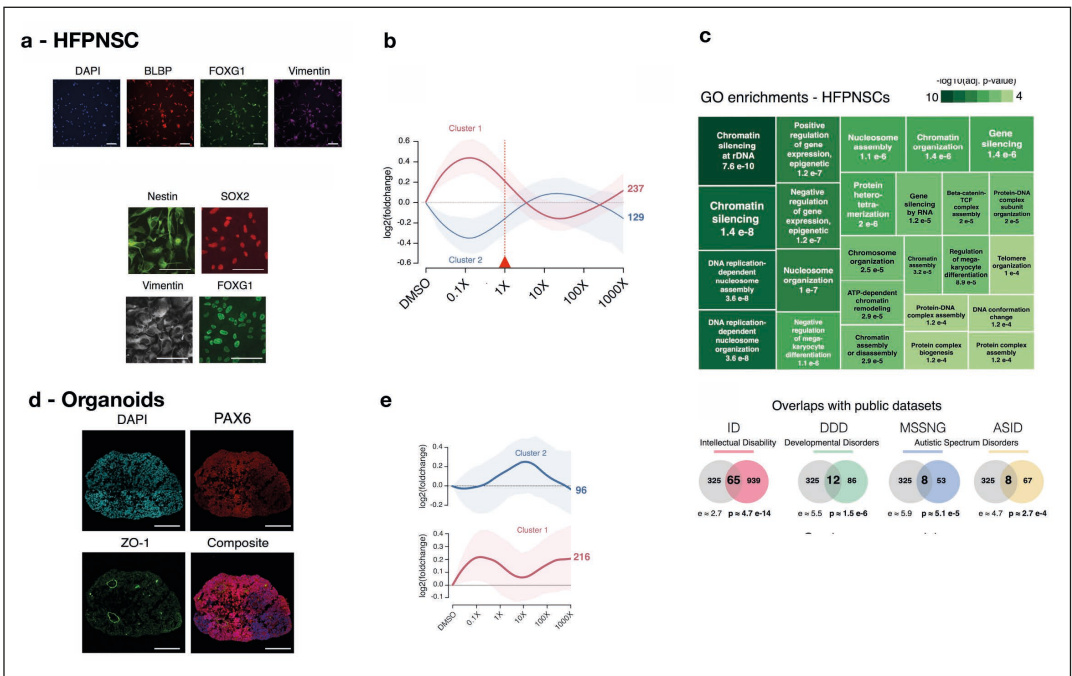


Fig. 2

(Figure 2j-k), we confirmed that in this different genetic background, MIX N yields non-monotonic dose-responses with dysregulation of gene expression at low concentrations and dysregulated genes that are significantly enriched for categories relevant for neurodevelopment, such as axonogenesis and synaptic signaling (Figure 2l-m).

Dissecting the impact of MIX N versus MIX G on human development

Despite having very similar chemical compositions, MIX N and MIX G had been linked to different health outcomes on the basis of epidemiological evidence. We thus sought the molecular basis of this distinction by cross-exposure of representative models to the alternative mixture. HFPNSC were exposed to MIX G, using the same five dilutions as for MIX N. As expected, the two mixtures showed a general transcriptomic impact of similar magnitude (369 unbiased DEGs for MIX N versus 275 for MIX G) and with significant overlap of the DEGs ($p \sim 1e-3$). Strikingly however, the mixtures showed marked differences in the affected genes, in particular with respect to ASD and ID associated targets that were only significantly enriched in MIX N DEGs, consistent with the association of MIX N exposure to early verbal skills. Moreover, even those DEGs associated with ASD or ID that were altered by both MIX N and MIX G were impacted at the lowest doses of MIX N but only at the highest doses of MIX G.

Together, these results provide experimental evidence of the mixture-to-phenotype dissection that had been originally only inferred at the population level, establishing the power of such integrated approaches for defining the molecular traces of EDC exposure across the population, organismal and cellular scales.

Discussion

The current vision for improving regulatory decision making relies on the transforming potential of high throughput and high content data to elucidate and quantify the molecular, cellular and organismal responses to chemicals (17). In the context of chemical regulation most authorities, including the Organisation for Economic Co-operation and Development (OECD), recommend integrated approaches to testing and assessment (IATA) that incorporate results from multiple methodologies. We first identify the adverse outcomes (language delay or low birth weight) in humans, then proceed to determine the chemical mixtures associated with these outcomes and, finally, establish the causative molecular and cellular impacts using *in vitro* and *in vivo* models. By making EDC mixtures experimentally tractable as the 'real life'-relevant unit of exposure, these complementary methodologies allowed us to uncover the gene networks specifically altered by neurodevelopment- or growth-targeting EDC mixtures and define thyroid function as a key and unifying axis of vulnerability to both mixtures. Furthermore, by establishing the value of human reprogrammed models based on self-renewing sources, we both expand their reach to regulatory toxicology and enrich the latter experimental human insight.

Together, this approach allowed us to define dysregulated gene networks, identified and validated through complementary methods, that can be broadly applicable within the new regulatory frameworks.

Bibliografia

1. Shi Y, Inou H, Wu, JCYamanaka S. Induced pluripotent stem cell technology: a decade of progress. *Nat. Rev. Drug Discov.* 2017; 16: 115-130.
2. Silbereis JC, Pochareddy S, Zhu Y, Li M, Sestan N. The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. *Neuron.* 2016; 89: 248-268.
3. Di Lullo E, Kriegstein AR. The use of brain organoids to investigate neural development and disease. *Nat. Rev. Neurosci.* 2017; 18: 573.
4. System BRF, Others S. Atlanta (GA): US Department of Health and Human Services, Centers for Disease Control and Prevention. National Center for Chronic Disease Prevention and Health Promotion. 2014.
5. Who U. State of the science of endocrine disrupting chemicals 2012: an assessment of the state of the science of endocrine disruptors prepared by a group of experts for the United Nations Environment Programme and World Health Organization. Geneva, WHO. 2013.
6. Kortenkamp A. Low dose mixture effects of endocrine disrupters and their implications for regulatory thresholds in chemical risk assessment. *Curr. Opin. Pharmacol.* 2014; 19: 105-111.
7. Anderson OS. et al. Epigenetic responses following maternal dietary exposure to physiologically relevant levels of bisphenol A. *Environ. Mol. Mutagen.* 2012; 53: 334-342.
8. Vandenberg LN, et al. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. *Endocr. Rev.* 2012; 33: 378-455.
9. De Rubeis, S. et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature.* 2014; 515: 209-215.
10. Vissers LELM, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. *Nat. Rev. Genet.* 2016; 17: 9-18.
11. Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. *Nature.* 2017; 542: 433-438.
12. Yuen CRK, et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat. Neurosci.* 2017; 20: 602-611.
13. Stessman HAF, et al. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat. Genet.* 2017; 49: 515-526.
14. Paşca AM, et al. Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nat. Methods.* 2015; 12: 671-678.
15. Dunaway KW, et al. Cumulative Impact of Polychlorinated Biphenyl and Large Chromosomal Duplications on DNA Methylation, Chromatin, and Expression of Autism Candidate Genes. *Cell Rep.* 2016; 17: 3035-3048.

16. Adamo A, et al. 7q11.23 dosage-dependent dysregulation in human pluripotent stem cells affects transcriptional programs in disease-relevant lineages. *Nat. Genet.* 2015; 47: 132-141.
17. Califf RM, et al. Transforming Evidence Generation to Support Health and Health Care Decisions. *N. Engl. J. Med.* 2016; 375: 2395-2400.

Non di solo DNA

Giuseppe Testa

European Institute of Oncology, Milano

Testo tratto da MicroMega 6/2017 – Almanacco della Scienza

‘Ce l’ha nel DNA!’: quante volte abbiamo sentito, e usato, questa frase per esprimere l’idea di una certa ineluttabilità di un carattere genetico? I più recenti studi di epigenetica mostrano però che le cose sono ben più complicate e che il genoma da solo è ‘muto’: per operare ha bisogno di informazioni ulteriori, che vengono influenzate anche dall’ambiente esterno. Un’ipotesi che apre scenari nuovi sulle nostre responsabilità nei confronti delle generazioni future.

Un Giano bifronte

Volendo seguire la prospettiva del *micro* e del *macro* che caratterizza questo Almanacco della scienza di *MicroMega*, potremmo dire che tra il *micro* e *macro* c’è l’*epi*. *Epi* come epigenetica, il campo delle scienze della vita, e della biomedicina più in generale, che studia a livello molecolare come ci portiamo il mondo dentro, come cioè l’ambiente e il tempo in cui viviamo plasmano il nostro DNA, attualizzandone il potenziale e curvandolo secondo pieghe di sviluppo e attività. Plasmare, curvare, piegare: tutte parole che ci parlano di plasticità, termine intrinsecamente, direi quasi fondativamente associato all’epigenetica, e che ne illumina una delle due facce. L’epigenetica infatti, non solo come disciplina ma proprio come risorsa discorsiva entrata a far parte, almeno nel mondo anglosassone, dell’immaginario collettivo di ampie fasce della popolazione, è un Giano bifronte, e anzi, come abbiamo di recente argomentato, è una disciplina che prospera proprio nell’ambiguità dei suoi significati (Meloni e Testa, 2014). Il nostro scopo dunque non è tanto quello di fornire una sintesi di questo campo scientifico, per la quale rimandiamo al documentario che abbiamo di recente realizzato nell’ambito del progetto bandiera sull’epigenetica EPIGEN del Cnr (disponibile sul sito del progetto al seguente link: goo.gl/ciWNTc), quanto quello di allertare il lettore, per spunti, sulle implicazioni concettuali di questo ambito della scienza e fornirgli gli strumenti critici per coglierne le promesse e le sfide.

La faccia 'plastica' dell'epigenetica...

Partiamo allora dalla sua prima faccia, attraverso cui l'epigenetica connota appunto la plasticità e la duttilità dei fenomeni vitali, ancorandole al DNA ma svincolandole dalla rigidità intrinseca all'idea del genoma come libro di istruzioni che esegue sé stesso, secondo la feconda definizione di Moss (Moss, 2003, p. 52). Se al genoma associamo la fissità di ciò che è scritto, l'epigenoma ci rimanda invece all'interpretazione di chi legge, ai molti modi con cui uno stesso testo può essere letto, interpretato e tramandato. Dei circa 25 mila geni delle nostre cellule, ogni tipo cellulare del nostro organismo ne utilizza solo alcuni, modulando quella che chiamiamo espressione genica volendo appunto alludere a un potenziale codificante che, per essere attualizzato, dev'essere innanzitutto espresso. Ed essere espresso significa, per un gene, venire utilizzato per sintetizzare, da parte della cellula, una proteina o un RNA (un tipo di acido nucleico analogo al DNA), i due elementi funzionali essenziali che determineranno poi l'identità della cellula, regolandone forma e funzione. Ecco, l'espressione genica copre appunto quel lungo e complesso flusso di operazioni che porta a leggere un segmento genico - di DNA - per farne un RNA o una proteina, e poi porta questi ultimi a svolgere la loro funzione attraverso un'ulteriore cascata di reazioni biochimiche che non sono scritte direttamente nel DNA ma sono invece plasmate dal contesto (cellulare, organismico, ambientale) che estrae via via, dal potenziale codificante di quel segmento di DNA, la sua funzione compiutamente attuata.

Eminente flessibilità - e anche reversibilità dunque - giacché quello che non è fisso, che può essere modulato dall'esterno è anche, quasi per definizione, reversibile. E questo certamente gioca un ruolo nel successo con cui il termine «epigenetico», e il mondo semantico che vi ruota attorno, si è insinuato tanto nella percezione quanto nell'immaginario collettivi. Termine coniato inizialmente dal biologo Waddington, che lo introdusse però, in un'ottica di sviluppo, per illuminare come nuovo campo di studio i processi che portano dal genotipo al fenotipo (Waddington, 2012), il mondo semantico dell'epigenetica si è poi espanso a un repertorio più ampio, e a sua volta molto duttile, di significati (Meloni e Testa, 2014; Boniolo e Testa, 2011). Nell'accezione ora comune risalta il prefisso *epi-*, dal greco per su o sopra, a denotare appunto tutto quello che sta sopra al DNA, quell'apparato di espressione genica brevemente richiamato prima, eminentemente malleabile agli stimoli ambientali nel senso più ampio del termine. Nella società liquida in cui le forme del nostro vivere sociale sono sempre più provvisorie, in cui tutto è pensato come sempre potenzialmente reversibile e riciclabile, in cui ciascuno è spinto ad assumere una molteplicità crescente di ruoli e in cui il tempo delle scelte è sempre più dilatato in una sequenza di opzioni, l'epigenetica fornisce dunque una serie sempre più vasta di dati, tecnologie e risorse discorsive che ne fanno una scienza compiutamente allineata alla contemporaneità.

Di questa faccia plastica l'esempio paradigmatico, e anche quello al momento più fecondo di implicazioni mediche, è la cosiddetta riprogrammazione cellulare, la capacità cioè di tramutare un qualsiasi tipo di cellula in un altro, praticamente *ad libitum*. Potrebbe sembrare di primo acchito sorprendente accostare programma a

plasticità, due parole che tendono verso polarità opposte nell'intuitiva dicotomia tra l'esecuzione di un programma e la plastica adattabilità della materia vivente. Ma l'epigenetica, almeno in questa accezione, le tiene assieme proprio perché ci svela la plasticità dei programmi di espressione genica. Decenni di biologia molecolare ci hanno portato infatti a pensare all'identità delle cellule - e a operare su di essa - come al risultato di un programma (quegli ordinati flussi biochimici cui abbiamo accennato sopra); programmi che però sono appunto plasticamente reversibili e modificabili perché non fissi nella sequenza del DNA ma tracciati in quel complesso apparato regolatorio «di lettura» che da quelle sequenze estrae appunto le varie forme e funzioni specializzate delle nostre cellule. Sulle riviste scientifiche più qualificate si parla ormai da anni esplicitamente di *cell fate plug and play* (Chambers e Studer, 2011), laddove l'identità cellulare viene appunto accostata al linguaggio del *plug and play* con cui eseguiamo una delle miriadi di operazioni digitali del nostro tempo, dall'ascolto di un brano musicale al caricamento di un video all'attivazione di un'app, qualcosa insomma che può essere appunto inserito (*plugged*) e poi eseguito (*played*) a mo' di programma. Il Nobel per la medicina del 2012 venne conferito a Gurdon e Yamanaka per aver, rispettivamente, aperto e compiuto, la parabola scientifica che ci consente di manipolare l'identità cellulare operando appunto sulla regolazione epigenetica della cellula (Colman, 2013). Ormai possiamo riprogrammare tutte le cellule del nostro corpo, incluse quelle di scarto come quelle che si sfaldano nelle nostre urine o quelle ottenute da un semplice prelievo di sangue, in cellule staminali pluripotenti, quel tipo cellulare presente solo all'inizio del nostro sviluppo, e solo transitoriamente. Cellule preziose perché capaci di generare, sia *in vitro* sia *in vivo*, tutti i tipi di tessuti. È questo che apre vaste frontiere biomediche, davvero impensabili fino solo a qualche anno fa. Perché, se da una cellula della mia pelle posso derivare un neurone o una cellula del muscolo cardiaco o addirittura degli abbozzi sempre più sofisticati di tessuti e organi (i cosiddetti organoidi), ecco che posso pensare di rigenerare parti del corpo che siano perfettamente compatibili dal punto di vista immunitario (perché derivate appunto dalle mie stesse cellule e recanti quindi il mio stesso genoma). È l'attualizzazione - ancora solo prospettica, almeno su scala - di una visione prometeica che già vede, in Giappone, la prima donna al mondo ad avere avuto l'impianto nell'occhio di un dischetto di epitelio pigmentato della retina epigeneticamente riprogrammato a partire dalla sua stessa cute (Mandai *et al.*, 2017). E non è solo la visione prometeica a essere in campo, ma anche quella, parallela ma complementare, dell'avatar, cui mi sono riferito ad esempio nei progetti del nostro laboratorio per descrivere i modelli di malattia, resi possibili tramite la riprogrammazione epigenetica, che ci stanno consentendo di studiare patologie come l'autismo e la disabilità intellettiva (Adamo *et al.*, 2015). Per la maggior parte delle malattie umane infatti, l'ostacolo finora maggiore in medicina è stata la scarsa disponibilità di cellule primarie malate perché queste erano appunto dentro al corpo dei pazienti. Ma nel momento in cui posso utilizzare la riprogrammazione epigenetica per esternalizzare tipi cellulari finora inaccessibili, sottraendoli alle profondità del corpo per renderli sperimentalmente trattabili in laboratorio, ecco che le malattie umane divengono - per la prima volta - osserva-

bili nel loro insorgere e nel loro divenire in queste rappresentazioni del corpo fuori dal corpo; avatar - ma in senso molto concreto e compiutamente carnale - delle malattie o anche semplicemente delle predisposizioni alla salute o alla malattia che ciascuno di noi si porta dentro.

L'oncologia è l'altro ambito in cui l'epigenetica sta acquisendo enorme importanza come frontiera di conoscenza e grande speranza di nuove terapie, una promessa che nasce, almeno in parte, dalla consapevolezza di una parziale sconfitta. Con la profilazione di genomi su scala sempre più vasta infatti, inclusi naturalmente i genomi dei tumori, si era molto sperato che la dettagliata conoscenza delle mutazioni genetiche permettesse di indirizzare al meglio le nuove terapie basate appunto sulla disponibilità dei cosiddetti farmaci molecolari, disegnati per colpire cellule tumorali recanti una specifica mutazione genetica e dunque una particolare alterazione proteica. Così sono partiti i cosiddetti «studi a cestino» (dall'inglese *basket trial*), in cui si è pensato di somministrare lo stesso farmaco a più pazienti non sulla base di una diagnosi istologica del tumore (quella che ha finora guidato la nostra architettura nosologica) bensì sulla base della comunanza di una specifica lesione molecolare: dimmi che genoma ha il tuo tumore e ti dirò di che farmaco hai bisogno, e quindi tumori di diversi tipi, in diversi organi, sono stati appunto accorpati nello stesso «cestino» sperando che lo sguardo molecolare si rivelasse più produttivo di quello finora regnante dell'anatomia patologica. Di quel regno Foucault ci ha tracciato mirabilmente la nascita, ricordandoci come «lo spazio della configurazione della malattia e lo spazio della localizzazione del male del corpo, non sono stati sovrapposti, nell'esperienza medica, che per un breve periodo: quello che coincide con la medicina del XIX secolo e coi privilegi accordati all'anatomia patologica. Epoca che segna la sovranità dello sguardo, poiché nello stesso campo percettivo, seguendo le stesse continuità o le stesse faglie, l'esperienza legge d'un sol colpo le lesioni visibili dell'organismo e la coerenza delle forme patologiche; il male s'articola esattamente sul corpo, e la sua distribuzione logica si fa d'acchito per masse anatomiche» (Foucault, 1998, pp. 15-16). Ecco, i *basket trials* rappresentano compiutamente la svolta epistemologica in cui quel periodo «tanto breve» si avvia esso stesso al termine, con la sostituzione dello sguardo molecolare allo sguardo anatomo-patologico, e con la distribuzione logica delle malattie che si fa non più, o quantomeno non solo, per masse anatomiche bensì per cataloghi di geni e delle loro lesioni. Il fatto è però che finora i *basket trials* sono falliti, almeno in larga parte (Hyman et al., 2015). Diversi tumori, pur avendo la stessa mutazione genetica, hanno risposto diversamente alla stessa terapia, riportando imperiosamente alla ribalta il contesto cellulare e rivalorizzando dunque lo sguardo «per masse anatomiche» ma traducendolo, naturalmente, nella sua declinazione contemporanea, in cui le masse anatomiche lasciano il posto all'identità delle cellule e alla loro impronta epigenetica, percepita come sempre più plastica e riprogrammabile, e quindi capace di rispondere alla stessa lesione genetica con diversi programmi di espressione a livello dei quali deve giocarsi la nuova sfida terapeutica. Da qui i titoli emblematici di due recenti articoli prospettici: «La plasticità epigenetica e le caratteristiche del cancro» (Flavahan, Gaskell, Bernstein, 2017) e «La tumorigenesi tessuto-specifica: il contesto conta» (Schneider et al., 2017).

... e la sua faccia 'stabile'

Essendoci ora familiarizzati con il lato plastico dell'epigenetica, veniamo all'altra faccia di Giano, perché l'epigenetica, almeno nella sua declinazione invece più rigorosa, si concentra su quelle interpretazioni di testo che sono durature, sui modi cioè in cui un qualsivoglia stimolo ambientale è sì in grado di curvare il potenziale informativo del nostro DNA, ma di farlo in maniera stabile. E la stabilità la possiamo ritrovare a vari livelli. Durante lo sviluppo, ad esempio, sono epigenetici quei meccanismi che plasmano via via l'identità delle nostre cellule, vale a dire la loro specificità di forma e funzione, senza che a ogni nuova divisione ci sia bisogno di ricominciare daccapo, quei meccanismi insomma che abbiamo visto sopra nel momento in cui dovevano essere smantellati per trasformare un tipo cellulare in un altro. Il vantaggio evolutivo di tali meccanismi sembra ovvio se pensiamo alla comparsa degli organismi multicellulari che si accrescono a partire da una singola cellula con la graduale generazione di cellule figlie, nipoti eccetera che acquisiscono via via le loro diverse identità o competenze. Pensiamo alla cellula iniziale che si duplica, e alle due cellule figlie che si duplicano a loro volta per arrivare ai milioni di cellule del nostro corpo distinte in tantissime forme e funzioni. Sappiamo dagli inizi degli anni Settanta, grazie agli studi di Gurdon evocato prima, che, almeno a grandi linee, questa progressiva specializzazione delle cellule non si accompagna a cambiamenti significativi nel genoma, e cioè che tutte le nostre cellule hanno, almeno in gran parte, lo stesso DNA. Ci diventa subito chiaro dunque che succederebbe se nel loro dividersi, le cellule madri non trasmettessero alle cellule figlie anche qualcos'altro rispetto al crudo DNA: non ci sarebbe sviluppo perché a ogni generazione di cellule si sarebbe cancellata la memoria dei passi già fatti dalla cellula madre in una determinata traiettoria (ad esempio nella progressiva formazione del sangue, del cervello, del muscolo eccetera). Non può essere solo dunque il libretto di istruzioni (cioè il crudo DNA) a passare di generazione in generazione; anzi, proprio perché il genoma non è un libretto di istruzioni che esegue se stesso, ma ha bisogno, per dispiegare il proprio potenziale istruttivo, di un apparato interpretativo che ne precisi via via il significato, ecco che, ogni qualvolta una cellula si divide, una parte di quell'interpretazione dev'essere essa pure trasmessa: quel tanto che basta per mantenere la traiettoria di sviluppo già intrapresa e al contempo restare plasticamente duttile per nuovi stimoli ambientali o esigenze adattative.

Fin qui abbiamo fatto riferimento all'epigenetica durante lo sviluppo, alla stabilità cioè di fenotipi cellulari che sono propagati più o meno fedelmente attraverso generazioni di cellule, a quei programmi che, come abbiamo visto nella prima parte, possono essere a loro volta resi reversibili proprio in quanto epigenetici. Ma la stabilità epigenetica può attraversare anche generazioni di organismi, ed è qui che naturalmente lo sguardo molecolare che si sta aprendo su questi meccanismi ci riporta all'ereditarietà dei caratteri acquisiti, con il suo controverso lignaggio percepito talvolta come compiutamente eretico, inserendo l'epigenetica molecolare in una moderna e multidimensionale lettura dell'evoluzione (Jablonka e Lamb, 2005). Perché fin quando sono le cellule di uno stesso organismo a tra-

smettersi, di generazione in generazione, l'impronta del mondo e del tempo in cui ciascuna è vissuta, questo è ancora pienamente compatibile con l'idea che di organismo in organismo sia solo il DNA a trasmettere i caratteri. Nella sua stentorea dicotomia, Weissman divideva appunto la linea somatica da quella germinale: da un lato dunque tutte le cellule del nostro soma, per le quali non è scandaloso ma anzi necessario pensare a meccanismi che gli assicurino la capacità di rispondere all'ambiente in maniera plastica ma al contempo stabile. La nostra memoria, i cambiamenti metabolici innescati da una dieta, la tonicità di un muscolo allenato sono altrettanti esempi di una regolazione genica, ormai potremmo dire di un programma, che diviene appunto stabile e viene mantenuto una volta innescato - almeno per un certo tempo - anche se l'innescato originale non è più presente. Dall'altro lato della demarcazione di Weissman c'è invece la linea germinale, cioè i gameti con cui generiamo le nuove generazioni. E qui la dicotomia era netta, con l'idea che il soma, pur registrando la varietà degli stimoli ambientali e portandosi dunque il mondo dentro, non trasmettesse alcuna di queste informazioni alla linea germinale, che passava dunque più o meno intonsa di generazione in generazione. Questa era dunque esposta solo alla mercé delle mutazioni genetiche (repertorio di variabilità) ma sottratta all'*hic et nunc* delle mutazioni epigenetiche, cioè di quei cambiamenti che interessassero non le sequenze del DNA ma i modi in cui quelle stesse sequenze potessero essere usate, meccanismi dunque capaci di agire su scale temporali infinitamente più brevi di quelle della selezione naturale operante sulla variabilità genetica.

Ormai da molti anni sappiamo però che in molte specie è invece possibile la cosiddetta ereditarietà epigenetica transgenerazionale. Vale a dire la trasmissione di fenotipi attraverso generazioni di individui, fenotipi che non sono legati a mutazioni nel DNA ma a stabili cambiamenti nella sua regolazione, innescati da stimoli ambientali in una data generazione e propagati alle generazioni successive senza più bisogno del ripetersi dello stimolo originale (Heard e Martienssen, 2014). Per ovvi motivi si tratta di fenomeni molto difficili da precisare nell'uomo - se non altro per la lentezza degli studi - e in cui le evidenze su base epidemiologica (legate ad esempio alle conseguenze di gravi carestie sulle generazioni successive) non hanno ancora consentito certezze definitive. Ma, oltre che nelle piante e in molti invertebrati, anche nei mammiferi, *in primis* nel topo, ci sono stati enormi passi avanti nella messa a punto di sistemi sperimentali in grado di capire se veramente un determinato tratto, acquisito in base a uno stimolo ambientale, potesse essere trasmesso alle generazioni seguenti in maniera indipendente dalla sequenza del DNA (Mansuy e Bohacek, 2017). Uno degli esempi recenti più strabilianti è stata l'ereditarietà di un fenotipo comportamentale causato dalla reazione di paura scaturita dall'aver percepito un particolare odore in associazione a una scarica elettrica. Non solo i figli, ma anche i nipoti e i pronipoti di topi sottoposti a questo trattamento reagivano con la stessa reazione di paura allo stesso odore, anche se non erano mai stati esposti all'associazione di odore e scarica elettrica (Dias e Ressler, 2014). Un risultato affascinante e inquietante al contempo, perché svela compiutamente l'altra faccia del Giano epigenetico, e cioè l'impatto della sua stabilità. Epigenetico non è solo quello che è plastico e duttile, ma quello che,

essendo plasticamente indotto, può essere poi mantenuto in modo duraturo, addirittura per generazioni. Ancora oggi lo scetticismo riservato a questi studi viene dal fatto che nei mammiferi i gameti, e l'embrione dopo la fecondazione, vanno incontro a un pervasivo processo di riprogrammazione epigenetica, che sappiamo cancellare la maggior parte dei marchi accumulatisi nelle cellule germinali durante la loro formazione. Da qui l'idea, ancora prevalente, che i gameti, e soprattutto lo sperma, agiscano davvero da *tabulae rasae*, passando al nuovo individuo il genoma e poco altro, e soprattutto un genoma liberato dalle impronte di qualsivoglia impatto ambientale possa averlo riguardato. Non poteva esserci insomma «odore» nello sperma, volendo riallacciarci all'esempio di sopra, e tantomeno una memoria dell'odore codificata nello sperma per via esclusivamente epigenetica. Almeno fino a poco tempo fa. Perché, di fronte a questi e molti altri dati che vanno nella stessa direzione, bisogna evidentemente ritenere che la riprogrammazione dei gameti non sia completa e che essi consentano, o addirittura si siano evoluti per promuovere, il passaggio transgenerazionale di alcune memorie epigenetiche, sotto forma magari di piccole molecole di RNA, come sembra emergere da numerosi studi.

Comunque, anche nella maggior parte di questi casi meglio caratterizzati, siamo ancora lontani dal chiarire i meccanismi molecolari che ne sono alla base, e l'epistemologia di questi studi procede ancora tutta per sottrazione. Avendo cioè definito nel genoma l'asse dell'ereditarietà, qualsiasi studio che voglia dimostrare la trasmissione transgenerazionale di un fenomeno per via solo epigenetica deve sempre poter dimostrare l'assenza di cambiamenti genetici che possano invece sottenderla.

Ereditarietà e responsabilità

Al di là della fascinazione per questi esperimenti, le loro implicazioni per il nostro sentire biografico non possono evidentemente sfuggire. Perché se quello che hanno mangiato o fatto o subito i nostri nonni o trisavoli può plasmare il modo in cui veniamo al mondo, sia pure solo in parte, si aprono scenari inediti, o quantomeno amplificati, di responsabilità individuali e collettive nei confronti del futuro (Chiapperino, G. Testa, 2016). In questo contesto è interessante l'ingresso del maschio sul proscenio della responsabilità genitoriale in era molecolare. Sono infatti soprattutto a carico dei maschi le prove più convincenti (ma anche quelle più veloci e facili da raccogliere) sulla trasmissione epigenetica transgenerazionale. Così il maschio, scudato per decenni dalla medicalizzazione rivolta alla donna, prima e dopo la gravidanza, in termini di responsabilità per la salute del nascituro, si trova all'improvviso catapultato in uno spazio non dissimile di presa in carico, in cui gli spermatozoi iniziano a essere visti e studiati come veicoli di salute o malattia che forse riflettono, almeno in parte, le esperienze dei padri. A proposito della politica della e sulla vita, Rose aveva perspicacemente osservato come la genetica, nella sua accezione più riduzionista che abbiamo già estensivamente analizzato e destrutturato (Nowotny e Testa, 2012), venisse appunto criticata per la sua «ontologia di profondità» (Rose, 2006, p. 130, tr. mia). I biologi «si ritiene che

costruiscano il codice genetico come una profonda verità interna, causa di salute o malattia, semplicemente espressa sulla superficie della corporalità, della condotta, del carattere eccetera» (*ibidem*). Eppure, dopo aver esaminato altre ontologie di profondità che caratterizzano il moderno, dall'economia politica alla psicoanalisi, Rose concludeva argomentando come «la genetica contemporanea cominci a operare in un mondo “appiattito”, un mondo di superfici più che di profondità. Negli schemi esplicativi della postgenomica, non si pensa più al codice genetico come a una struttura profonda che causa o determina, ma come a un set di regole in network, filiazioni e connessioni complesse, ramificate e non gerarchiche» (*ibidem*). Al di là delle valutazioni su quest'analisi della postgenomica, che coglie l'epistemologia per superfici che caratterizza le scienze della vita contemporanee ma stride con la dirompenza mai tanto ambiziosa ed efficace della nuova genomica umana su larga scala, è interessante cogliere l'andamento apparentemente antitetico di genomica ed epigenomica. Nel senso che proprio nel momento in cui il gene sembra, almeno ad alcuni, recedere dal suo primato ontologico di profondità, anche grazie alla rinnovata attenzione per l'epigenetica, questa a sua volta si candida a prenderne il posto, nella sua componente transgenerazionale, allineando i presupposti per un'ontologia di profondità tanto più rilevante quanto maggiore è, in una biografia umana, la ricchezza di esperienze potenzialmente trasmissibili.

Proprio con il biografico e l'umano possiamo dunque chiudere queste riflessioni, non senza però un ultimo passaggio al moscerino della frutta, la *Drosophila melanogaster* servita da modello chiave per lo sviluppo della genetica e della biologia molecolare. Proprio nel moscerino infatti si è appena chiarito il meccanismo molecolare di un fenomeno epigenetico transgenerazionale, legato alla plastica riorganizzazione della cromatina, la complessa struttura di proteine e RNA entro cui il genoma viene appunto letto e interpretato dalla cellula all'interno del nucleo (Ciabrelli et al., 2017). Modifiche di questo assetto sono risultate essere appunto trasmissibili attraverso varie generazioni di moscerini, anche in assenza della modifica sperimentale originale, che era inizialmente fornita artificialmente ma che è poi stata testata anche nella simulazione di un «vero» stimolo ambientale, e cioè con la variazione della temperatura. Gli autori scrivono infatti di aver «esplorato la stabilità delle epilinee [le linee di moscerino epigeneticamente modificate] in condizioni rilevanti dal punto di vista ecologico, ricreando un ambiente naturale in microcosmi». Ecco, l'evocazione di questi microcosmi sperimentali chiude il cerchio con il *micro* e il *macro* da cui siamo partiti, con l'*epi* di epigenetica in mezzo non a significare una dimensione intermedia - naturalmente - giacché sempre di meccanismi molecolari, e dunque *micro*, si tratta. Ma a significare appunto l'ambizione e la crescente capacità di un campo di ricerca di interpersi, come interfaccia epistemologica, tra il *macro*, cioè l'ambiente nella sua vastità spaziotemporale, e il *micro* inteso come dinamica molecolare delle nostre cellule e dei nostri tessuti, in cui leggiamo il *macro* dopo averlo tradotto, epigeneticamente, in impronte *micro*. Come abbiamo infatti di recente argomentato, «la promessa dell'epigenetica è quella di catturare la vastità analogica dei segnali ambientali [...] attraverso la rappresentazione digitale delle risposte molecolari che questi inducono. Se quello che sembrava irriducibilmente analogico (il sociale, l'am-

bientale, il biografico, l'idiosincraticamente umano) dev'essere disteso sul genoma compiutamente informatizzato dell'era digitale, in un flusso di reciproca reattività, ciò sembra essere possibile solo nel momento in cui quell'analogico viene interrogato, analizzato e rielaborato in rappresentazioni digitali compatibili con il genoma e con la nozione stessa di codice» (Meloni e Testa, 2014, p. 5).

Bibliografia

1. Adamo A, et al. 7q11.23 dosage-dependent dysregulation in human pluripotent stem cells affects transcriptional programs in disease-relevant lineages. *Nature Genetics*. 2015; 47: 132-141.
2. Boniolo G, Testa G. The Identity of Living Beings, Epigenetics, and the Modesty of Philosophy. *Erkenntnis*. 2011.
3. Ciabrelli F, et al. 2017. Stable Polycomb-Dependent Transgenerational Inheritance of Chromatin States in *Drosophila*. *Nature Genetics*. 2011 49: 876-886.
4. Chambers SM, Studer L. Cell Fate Plug and Play: Direct Reprogramming and Induced Pluripotency. *Cell*. 2011; 145: 827-830.
5. Chiapperino L, Testa G. The Epigenomic Self in Personalized Medicine: Between Responsibility and Empowerment», in Meloni M, Williams S, Martin P. (a cura di), *Biosocial Matters: Rethinking Sociology-Biology Relations in the Twenty-First Century*, Wiley-Blackwell, Hoboken. New Jersey. 2016.
6. Colman A. Profile of John Gurdon and Shinya Yamanaka, 2012 Nobel Laureates in Medicine or Physiology. *PNAS*. 2013; 110: 5740-5741.
7. Dias BG, Ressler KJ. Parental Olfactory Experience Influences Behavior and Neural Structure in Subsequent Generations *Nature Neuroscience*. 2014; 17: 89-96.
8. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic Plasticity and the Hallmarks of Cancer. *Science*. 2017; 357: 6348.
9. Foucault M. *Nascita della clinica. Una archeologia dello sguardo medico*, Einaudi, Torino. 1998.
10. Heard E, Martienssen RA. Transgenerational Epigenetic Inheritance: Myths and Mechanisms. *Cell*. 2014; 157: 95-109.
11. Hyman D., et al. Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. *The New England Journal of Medicine*. 2015; 373: 726-736.
12. Jablonka E, Lamb MJ. *Evolution in Four Dimensions. Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. MIT Press, Cambridge MA. 2005.
13. Mandai M, et al. 2017. Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration. *The New England Journal of Medicine*. 2005; 376: 1038-1046.
14. Mansuy IM, Bohacek J. A Guide to Designing Germline-Dependent Epigenetic Inheritance Experiments in Mammals. *Nature Methods*. 2017; 3: 243-249.
15. Meloni M, Testa G. *Scrutinizing the Epigenetics Revolution*. Biosocieties. 2014.
16. Moss L. *What Genes Can't Do*, MIT Press, Cambridge MA. 2003.

17. Nowotny H, Testa G. *Geni a nudo*, Codice edizioni, Torino. 2012.
18. Rose N. *The Politics of Life Itself: Biomedicine, Power, and Subjectivity in the Twenty-First Century*, Princeton University Press, Princeton. 2006.
19. Schneider G, Schmidt-Supprian M, Rad R, Saur D. Tissue-Specific Tumorigenesis: Context Matters. *Nature Reviews Cancer*. 2017; 17: 239-253.
20. Waddington CH. The Epigenotype. *International Journal of Epidemiology*, 2012; 41: 10-13.

