

36
PICCOLA
BIBLIOTHIKI

LEGGERE LA COMPLESSITÀ

Epigenetica, il DNA che impara

Ernesto Di Mauro

Epigenetica

il DNA che impara

*Istruzioni per l'uso
del patrimonio genetico*



Asterios Editore

Trieste 2017

Prima edizione nella collana PB: Gennaio 2017
©Ernesto Di Mauro, 2016
©Asterios Abiblio editore 2016
posta: asterios.editore@asterios.it
www.asterios.it

I diritti di memorizzazione elettronica,
di riproduzione e di adattamento totale o parziale
con qualsiasi mezzo sono riservati.
Stampato in UE.

ISBN: 978-88-9313-048-6

Indice

Introduzione, 9

CAPITOLO I

Il DNA

1. Siamo il prodotto di informazione, 12
*I termini del problema:
Fenotipo e Genotipo, Ploidia, Asimmetria.
Variabilità, Manovrabilità, Complessità.*
2. DNA e Progetti Genoma, 21
*Decifrazione del materiale genetico. Il concetto di "Testo"
riferito al DNA. L'unità del Testo. L'avvenuta decifrazione.*
3. Dall'informazione all'uso, 31
*Funzioni semplici e funzioni complesse.
Genomica comparata. Altri Genomi.
Single nucleotide polymorphisms (SNPs).*
4. Faber fortunae suae: Genetica ed Epigenetica, 42
5. L'infinito DNA, 44

CAPITOLO II

EPIGENETICA: ESEMPI

1. Il DNA è la nostra vera struttura, 47
2. Imago mundi, 48
3. Esempi estremi del ruolo dell'Epigenetica, 50
*Tempi lunghi: scomparsa e ricomparsa delle ali.
Tempi brevi: zigoti.
Tempi brevissimi: regolazione genica.*
4. Epigenetica come controllo della complessità, 57

CAPITOLO III

EPIGENETICA: I MECCANISMI

1. Epigenetica all'opera, 59
2. Interazioni evolutive tra Genetica ed Epigenetica, 64
3. Libero arbitrio e via di fuga, 67
4. Le ombre del passato nel nostro DNA, 67
5. Trasmissioni transgenerazionali, 68

6. Trasmissione transgenerazionale dello stress, 69
7. Epigenetica del sistema ipotalamico-pituitario-adrenalinico (HPA), 72
 8. A grandi linee, 75

CAPITOLO IV

EPIGENETICA DELLA PECORA DOLLY

1. Programmazione e deprogrammazione cellulare, 77
 2. Una cellula staminale è un embrione?, 82
 3. Crescita di nuovi arti, 83
 4. Dolly, 85
 5. Vivere bene e a lungo, 92
 6. CRISPR, 96

CAPITOLO V

ASPETTI ETICI DELL' EPIGENETICA: QUALCHE DATO PER POTER AFFRONTARE IL PROBLEMA

1. Elementi di storia naturale per riflettere sugli Organismi Geneticamente Modificati, 99
2. Soluzioni offerte dalla natura: il cordone ombelicale, 105
 3. Soluzioni offerte dall'ingegneria genetica, 106
 4. Contratto sociale, culturale, biologico, 108
 5. Prudenza e moderazione, 111

Conclusioni, 113

Indice delle finestre

1. WATSON E CRICK, 13
2. TOTIPOTENZA, 19
3. ANALISI DI ESPRESSIONE GENICA A LIVELLO GLOBALE, 30
 4. TECNOLOGIE DI BASE:
LA COSTRUZIONE DI DNA GENE CHIPS
ED IL MICROARRAYS DISPLAY, 41
 5. SILENZIAMENTO GENICO, 53
 6. CELLULE STAMINALI, 79
 7. CELLULE STAMINALI EMBRIONALI (ESCs)
E GERMINALI (GSCs), 80
 8. ANIMALE TRANSGENICO, 85
9. L'INVECCHIAMENTO È UN FENOMENO REGOLATIVO, 94
10. IDENTIFICAZIONE E PRODUZIONE DI GENI, 98

Introduzione

Questo libro parla di Epigenetica. Con il che si intende, formalmente: *la trasmissione di tratti e comportamenti senza cambiamenti della sequenza genica*. Il filo centrale di questo discorso è l'insieme di meccanismi che permettono e determinano l'uso e la trasmissione del patrimonio genetico, dagli aspetti più meccanicamente organizzativi a quelli funzionali, fino a quelli apparentemente più astratti, come la trasmissione del comportamento e della cultura. Poiché però il concetto stesso di Epigenetica non è precisabile se non abbiamo ben chiaro cosa si intende per Genetica, va prima definito cosa voglia dire questa parola oggi. Mi rivolgo dunque inizialmente a chi non sa bene cosa significhi Genetica nè, a maggior ragione, cosa sia l'Epigenetica; pur vivendone, pur essendo, come tutti noi, la loro incarnazione.

È importante capire cosa siamo, come siamo organizzati, quali siano le istruzioni per l'uso del genoma che abbiamo ricevuto. È importante rendersi conto che nasciamo con un hardware genetico di DNA privo di istruzioni per l'uso; e che, vivendo, il DNA crea il software per il proprio funzionamento. Il DNA impara semplicemente, vivendo.

Infine cercherò di dare corpo alle indicazioni, se proprio non vogliamo considerarle prove, che il comportamento e la cultura (nel senso più coinvolgente, più biologico della parola) si trasmettono per via epigenetica; e

non solo con esempi, immagini o parole. La trasmissione della cultura è un importante sistema evolutivo. Secondo quali regole? Esiste una Epigenetica del comportamento e, in senso più ampio, della cultura? Il DNA impara.

CAPITOLO I

Il DNA

1. Siamo il prodotto di informazione. *I termini del problema: Fenotipo e Genotipo, Ploidia, Asimmetria. Variabilità, Manovrabilità, Complessità.*
2. DNA e Progetti Genoma. *Decifrazione del materiale genetico. Il concetto di "Testo" riferito al DNA. L'unità del Testo. L'avvenuta decifrazione.*
3. Dall'informazione all'uso. *Funzioni semplici e funzioni complesse. Genomica comparata. Altri Genomi. Single nucleotide polymorphisms (SNPs).*
4. Faber fortunae suae: Genetica ed Epigenetica.
5. L'infinito DNA

Il mondo biologico è regolato dalle regole ferree della codificazione dell'informazione, della competizione e dell'evoluzione. La nostra cultura ha spesso difficoltà ad accettare la durezza di queste regole. Un approccio a questo aspetto della realtà che può in qualche modo mitigarlo consiste nel rendersi conto delle ragioni della sua ineluttabilità. Le pagine che seguono espongono brevemente i termini del sistema, cercando di fornire indicazioni su cosa si intenda oggi con le parole Genetica ed Epigenetica e descrivendo la storia recente della decifrazione dei Genomi.

Il DNA

Poche discipline hanno avuto uno sviluppo rapido come quello della biologia contemporanea. La conseguenza è che si è compresa quasi all'improvviso la natura del materiale genetico, il DNA, che determina l'essenza degli organismi viventi. Ne conosciamo la struttura, l'organiz-

zazione, i meccanismi di autoriproduzione; ne abbiamo compresa la qualità essenziale: quella di essere allo stesso tempo materia chimica inerte e materia vivente (il DNA si riproduce, si accresce, muore). La rapidità di queste scoperte ci ha impedito di acquisire fino in fondo consapevolezza che la natura degli organismi (e quella dell'organismo che ci interessa di più, *Homo sapiens*) è natura autocodificata, autogenita, autoreplicante. La lezione di libertà esistenziale che ne deriva, lungi dall'essere diventata patrimonio della cultura generale, non è ancora filtrata attraverso una diffusa coscienza. Entrambi i fatti (la autocodificazione e la mancanza di una coscienza di questa autocodificazione) meritano una riflessione. Merita anche riflessione il fatto che, una volta decifrato il codice e capito come questo codice è fisicamente organizzato, *Homo sapiens* ha immediatamente iniziato a manipolarlo. Guardandoli con un minimo di prospettiva, i dubbi, le moratorie, le proibizioni sembrano destinati a durare poco.

Il DNA si sta però rivelando un sistema molto più complesso di quello che sembrava: non è solo l'hardware del sistema biologico, ne è al tempo stesso il software modificabile e continuamente modificato. Il sistema è molto più delicato, elegante, efficace, complesso e vivo di quanto sembri.

1. Siamo il prodotto di informazione

Una struttura vivente è il prodotto di informazione. Nei nostri cromosomi centinaia di milioni di unità di informazione (i nucleotidi) sono messi in fila in ordine perfettamente (geneticamente) definito.

Un computer procede con scelte successive binarie, si-si-no-si-no-no... con informazioni date con due lettere sole, 0-0-1-0-“1”-1... . Il DNA è organizzato nello stesso modo: le sue cifre di base sono due (un nucleotide fatto

di purina ed uno di pirimidina), allineate lungo un filamento. Per aumentare la propria ricchezza di informazione il codice binario del DNA a un certo momento della propria evoluzione si è sdoppiato e sia le purine che le pirimidine sono diventate due: A e G sono purine, C e T sono pirimidine. Il filamento inoltre si avvolge e si sostiene intorno ad un filamento speculare. Per questo il DNA è una doppia elica.

Queste complessità aggiuntive non modificano la natura semplice dell'informazione genetica, chip lunghissimo fatto di un filo lineare in grado di autoriprodursi. Ogni filamento è stampo dell'altro, ogni filamento è quindi copia e stampo di se stesso.

Il numero di combinazioni possibili e diverse delle centinaia di milioni di unità di informazione che ci compon-

Finestra 1. Watson e Crick

James D. Watson (1928) e Francis H. Crick (1916-2004) hanno identificato la struttura del DNA nel 1952, in gran parte basandosi su dati cristallografici raccolti da Rosalyn Franklin (1920-1958). La descrizione della scoperta è stata pubblicata sul numero di *Nature* del 25 aprile 1953 (vol 4356, pag. 737) in un articolo dal titolo "Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid" ed ha valso loro il Nobel per la Fisiologia e la Medicina del 1962, condiviso con Maurice Wilkins. La frase finale dell'articolo è straordinaria: "Non è sfuggito alla nostra attenzione il fatto che lo specifico appaiamento che abbiamo postulato suggerisce immediatamente un possibile meccanismo di copiatura del materiale genetico". Questo concetto ha cambiato la storia della biologia e della scienza moderna. La struttura ipotizzata è basata sulla complementarietà chimica delle basi nucleiche che con la loro interazione assicurano stabilità alla struttura e sulla organizzazione a doppia elica speculare del filamento molecolare, a formare una struttura intrinsecamente riproducibile.

gono è inimmaginabile. Poiché siamo, siamo frutto di scelta, di selezione. Siamo informazione. Secondo la Scienza dell'Informazione questo significa in termini formali che siamo il prodotto di accumulo di energia.

Poiché discendo dai miei genitori, e i miei nonni erano già un po' più vicini a Neandertal, e questo a uno scimmione figlio di un lemure figlio di un topo figlio di uccelli che non sapevano più volare e così via per tre quattro miliardi d'anni..., poiché siamo informazione che non si è mai per definizione interrotta, siamo il frutto di un accumulo d'energia enorme. Siamo eredi di scelte evolutive fatte in tempi lontani e siamo depositari d'energia.

L'universo, qualsiasi sia la sua origine, sembra tendere alla disorganizzazione. Aumenta, come si dice, la sua entropia. In termini biologici, entropia significa morte. Contrario dell'entropia è l'organizzazione, la vita, l'informazione. Organizzazione è trattenerne energia in una struttura definita, finché questa non viene dispersa nei suoi frammenti componenti; vita è impedire la disgregazione.

Il lunghissimo processo evolutivo che ha portato alla identità genetica di ognuno di noi è partito all'inizio del mondo, per caso e per necessità. Ad ogni generazione alcuni organismi vengono biologicamente contro-selezionati (non si riproducono, o si riproducono meno degli altri, o muoiono prima). Ad ogni generazione, da miliardi di anni, si fanno delle scelte. E tutti gli organismi (la cui nascita e la cui sopravvivenza ha richiesto energia) contribuiscono a questa selezione, a questa lotta alla disorganizzazione, a questo formarsi e tramandarsi di informazioni. Ma come si esprime questa informazione, questo accumulo di energia? Oggi sappiamo che si esprime attraverso il DNA.

I termini del problema

Il nostro corpo è formato da cellule molto diverse tra loro. La diversità e la complessità che ne conseguono

sono innegabili. Per poterne capire il significato bisogna considerare tre fatti biologici fondamentali:

Fenotipo e genotipo. Queste categorie biologiche sono di facile definizione; nel loro insieme racchiudono tutte le strutture viventi. Capire fino in fondo cosa in realtà esse significhino è però più complesso. Il fenotipo è quanto di un organismo possiamo vedere, quello che funziona, la proteina muscolare che si contrae, il neurone che trasporta o elabora l'impulso nervoso, la struttura salina e cellulare dell'osso. Il nostro corpo è il nostro fenotipo, ciò che di noi stessi appare. Fenotipo: epifania, l'apparizione. Fanerogame, le piante che mostrano il proprio sesso, il proprio fiore. Fenomeno, ciò che è osservabile. Fenologia, le manifestazioni che animali e vegetali presentano nel corso di un anno.

La combinazione degli elementi, il filo che abbiamo nelle cellule, è il genotipo, il segreto riposto, testamento tramandato. Genotipo è ciò che non appare ma che determina. Genotipo: generare, genealogia, generazione, genere, genesi, gene. Genotipo è lo scritto dal quale si traggono le istruzioni per il fenotipo, è il libro che va letto, il testamento di segni che non va manomesso. Il genotipo è dunque il DNA; è la molecola che, rimanendo nell'ambito degli esempi appena fatti, fornisce l'informazione per sintetizzare le proteine muscolari, per strutturare i neuroni, per guidare le vie metaboliche che organizzano i depositi minerali di cui è fatto lo scheletro che ci sostiene. Inoltre, il DNA non è soltanto informazione pura. Essendo una molecola chimica ben definita e perfettamente organizzata in atomi e legami come le altre molecole, informative e non, il DNA è anche fenotipo. Il DNA è al tempo stesso informazione (nel senso di potenzialità attuali e virtuali, di trasmissione, di codificazione, di programmazione) ed è struttura fisica, è genotipo e fenotipo contemporaneamente.

La difficoltà di fondo di questo discorso è costituita

dalla impossibilità di separare rigorosamente i due concetti, disgiungere il corpo dalla informazione che lo determina, distinguere tra potenzialità e realtà. Come a lungo si è cercato di separare il corpo dall'anima, lo pneuma dal logos, la mente dal cervello, per poi capire l'inutilità dello sforzo e l'errore di fondo che questi tentativi costituivano, così è sbagliato separare genotipo da fenotipo. Le due definizioni sono complementari e speculari, volti didascalicamente separati di una realtà unica e complessa.

L'evoluzione si esercita *sul* fenotipo, sul muscolo della gazzella per conferirle lo scatto che la salva durante la fuga, sul cervello che ne coordina i movimenti. Ma si esercita *attraverso* il DNA, attraverso l'informazione che può trasferire alla successiva generazione di gazzelle uno spettro più ampio di possibilità di scatto e di fuga.

L'ambiente e la selezione sceglieranno se vale la pena, se conviene investire ulteriori energie fisiologiche in questo aspetto specifico della struttura, o se non convenga piuttosto affinare ad esempio la vista e l'udito. La scelta è combinatoria, passa attraverso il sistema muscolare, il corpo nel suo insieme, il comportamento, il DNA e la sua replicazione, la sua capacità di mutare lungo il filo delle generazioni e rimanere comunque DNA di gazzella. La combinatoria della evoluzione passa attraverso le scelte esercitate dall'ambiente nello scorrere degli anni, sovrapponendosi alla evoluzione dei predatori, cerchi concentrici di cambiamenti generati sulla superficie dello stagno del presente da sassi gettati dalla mano del tempo che passa. I cerchi si incontrano, si sovrappongono, interferiscono, in un rapporto evolutivo di equilibrio. La gazzella migliore sopravvive. La gazzella il cui genotipo è migliore e si incarna nel fenotipo migliore, genotipo e fenotipo *migliori* perché più adatti a quell'ambiente in quel momento, equilibrio evolutivo, frutto di miliardi di anni di selezione di strutture viventi che hanno avuto successo per il fatto stesso di essersi riprodotte all'interno del

gioco di competizione delle forme possibili, sulla superficie di questo pianeta, che è l'unico che abbiamo.

Ploidia. L'informazione genetica, il genotipo, è uguale in tutte le cellule somatiche, cellule il cui corredo genetico *diploide* (due copie) proviene in ognuno di noi da due genitori in parti sostanzialmente uguali. Il soma di un individuo ospita un gruppo di cellule il cui numero è relativamente limitato e che si comportano in modo funzionalmente specializzato nella loro finalità genetica: le cellule germinali. Le cellule germinali mature, sia maschili che femminili, hanno un corredo genetico *aploide* (una copia). Il processo differenziativo che porta al loro sviluppo e alla loro maturazione ne dimezza la ploidia e le programma in modo che fondendosi a coppie, una maschile e una femminile, si ricrei la diploidia originaria nella nuova cellula derivata da fecondazione (nello zigote cioè, fusione di un oocita aploide e di uno sperma aploide).

Lo zigote formatosi con la fecondazione comincia subito a dividersi e produce due cellule, e da queste quattro, otto e così via. Il processo è all'inizio simmetrico e segue regole numeriche semplici; poi, man mano, si complica. Nella natura stessa del processo è insita una tendenza alla asimmetria: (a) l'insieme aploide dei cromosomi (23 nella specie umana) è *quasi* uguale tra uomo e donna: 22 sono uguali, uno (X o Y) è differente. (b) Il contenuto di DNA è *quasi* uguale nell'oocita e nello sperma. Nell'oocita c'è qualcosa in più: i mitocondri, organelli con un loro DNA che si trasmettono esclusivamente per via materna. (c) La cellula uovo è *quasi* sferica e *quasi* simmetrica. Già dall' '800 si era cominciato ad osservare l'esistenza di gradienti di polarità, di *asimmetrie* sottili e complesse.

Asimmetria. La asimmetria più importante nei sistemi biologici è quella che determina la dinamica dello sviluppo. La cellula fecondata è totipotente, è una cellula sta-

minale totale, nel suo genoma ridiventato diploide è racchiusa tutta l'informazione per tutto l'organismo in tutti i suoi momenti di vita, embrionali, fetali, giovanili, adulti. Nella cellula appena fecondata sono iscritte le istruzioni per il bilancio metabolico che caratterizzerà la nostra vecchiaia. Il sistema è ovviamente di enorme complessità, sia perché deve rispondere ad un ambiente (esterno o interno) potenzialmente molto variabile, sia perché i programmi di attuazione sono dinamici e differenziali. Molto presto nello sviluppo le cellule si specializzano e perdono la propria potenzialità generale. Specializzazione e perdita di potenzialità sono due aspetti speculari dello stesso fenomeno. Da un punto di vista informazionale il processo è ovvio: se non ci fosse differenziamento, saremmo un unico grande ammasso di cellule uguali allo zigote, una enorme *morula*, usando un termine della embriologia classica. Da un punto di vista reale la perdita di potenzialità e l'attuazione differenziale dei programmi di specializzazione è un processo molto complesso. Uno degli aspetti più importanti di questa complessità è la sua stabilità differenziale: fino a che punto una cellula differenziata rimane tale? Come, quando e quanto si differenzia? Come, quando e quanto questo processo è controllabile e manipolabile, o spontaneamente reversibile?

Ci siamo così avvicinati ad uno dei problemi centrali della trasformazione tumorale delle cellule, sostanzialmente causata dalla perdita del sistema di controllo della loro riproduzione; il blocco della replicazione è tipico delle cellule differenziate che in termini generali si replicano meno di quelle che non lo sono, è tipico delle cellule differenziate ma è vero anche per quelle che lo sono meno. Ed è questo il punto centrale della ricerca biomedica contemporanea: è possibile usare in modo controllato i processi di differenziamento/sdifferenziamento per ottenere strutture biologiche per trapianti, sostituzioni, ricrescite, riparazioni di parti della macchina umana? È

possibile cioè osare quello che la natura normalmente non fa: passare al di là dei limiti posti statisticamente alla durata ed alla qualità della nostra vita.

Variabilità, manovrabilità, complessità. Così come sono stati definiti, questi processi di creazione di asimmetria sono, nei molteplici sistemi biologici, in grado di variare ed evolvere. Il sistema vivente è nel suo insieme estremamente plastico. Definita una regola è sempre possibile trovare in un trattato di zoologia o di botanica la sua eccezione.

Per capire meglio la portata di queste affermazioni sono importanti tre considerazioni: i rapporti genotipo/fenotipo, i rapporti soma/linea germinale, il rapporto diploide/aploide.

Tutto è manipolabile: tutti quelli che venivano considerati dogmi in embriologia, o istologia, o genetica che fosse sono caduti, uno dopo l'altro. Il cosiddetto dogma

Finestra 2. Totipotenza

La cellula appena fecondata, lo zigote, è totipotente. Man mano che la replicazione cellulare procede, le cellule perdono potenzialità generica ed attuano programmi differenziativi. Il momento di partenza di una linea cellulare specializzata è una cellula specifica, il capostipite in grado di dividersi e produrre discendenti uguali tra loro in quantità controllata. Queste cellule capostipiti, le cellule staminali, sono importantissime. Su di esse poggia la possibilità più promettente di affrontare in modo sostanziale tutte quelle malattie che comportano degenerazioni irreversibili di tipi cellulari definiti (e a volte, per ragioni di cooperazione negativa, di interi tessuti). I casi più noti sono quelli che riguardano le degenerazioni delle cellule delle isole del pancreas in molte forme di diabete, e le degenerazioni di neuroni nelle malattie di Parkinson, di Huntington, in molte altre condizioni neurologiche.

della biologia molecolare: “*il DNA fa RNA e questo fa proteine*”, è rimasto un dogma per pochissimi anni, travolto da una valanga di eccezioni; il dogma che le cellule nervose differenziate non si possono riprodurre sta facendo la stessa fine; la necessità dello sperma come principio di attivazione della replicazione iniziale è concetto dimenticato; e così via. Il limite è come al solito quello posto dalla conoscenza e dal suo progredire.

Il sistema va comunque sempre considerato in ambito statistico. Le regole alla base del funzionamento dei meccanismi genetici di genomi grandi come quello umano si avvicinano molto alle regole che sono alla base del modo di essere dei sistemi complessi. E, anche se in grado minore, lo sono anche le regole che si osservano in genomi relativamente più semplici quali quelli batterici. Ci riferiamo qui alla complessità che rispecchia la molteplicità di scelte contemporanee, ognuna deterministica se presa singolarmente, il cui risultato causa un insieme di scelte successive altrettanto complesse. Dato l'altissimo numero di reazioni contemporanee in un organismo vivente, la natura termodinamica delle reazioni, gli aspetti cinetici degli equilibri apparenti raggiunti, le interazioni con ambienti altrettanto complessi in continua modificazione, date tutte queste variabili non meraviglia che genomi uguali in condizioni simili rispondano in modo apparentemente diverso. Questo sostanzialmente significa che quando un virus infetta una cellula, il suo acido nucleico può entrare, cominciare a replicarsi e poi fermarsi e morire, senza conseguenze. Oppure può uccidere o trasformare la cellula vittima. Data la complessità e la natura stocastica dei molti processi coinvolti, definire la differenza causale tra i due comportamenti è sostanzialmente impossibile, è solo possibile capire gli schemi di riferimento quantitativi del fenomeno. Le conseguenze per l'individuo specifico sono però tutto-nulla, vita-morte.

Ed anche: quando diciamo che un virus particolarmente letale uccide il 93 per cento degli individui colpiti, che

spiegazione ci diamo della sopravvivenza del rimanente 7 per cento?

Ed ancora: se stabiliamo che nella specie umana in certe condizioni genetiche il tasso di mutazione è di 10^{-7} , questo significa che il DNA delle mie cellule si rompe in modo spontaneo con una certa frequenza e con una certa definizione di posizione. Dato questo valore numerico, capire dove, come, quando, e che conseguenze questo comporta per ogni singolo organismo nella sua individualità è naturalmente statisticamente complesso ma è anche, almeno per me se il DNA in gioco è quello mio, singolarmente rilevante.

2. DNA e Progetti Genoma

Il nostro materiale genetico, il DNA, è composto dalla ripetizione di quattro sostanze chimiche definite. Questi componenti, i nucleotidi, sono in parte uguali tra loro, in parte diversi. Le parti diverse sono costituite di molecole specifiche (l'Adenina, la Timina, la Guanina e la Citosina). I nucleotidi sono legati uno all'altro a formare due fili lunghissimi intrecciati tra loro: la doppia elica. Il filo dell'uomo è lungo più di quattro miliardi di lettere, siano esse A, T, G o C (indicandole con gli acronimi dei nomi delle sostanze chimiche reali). Queste lettere formano parole, racconti, lunghe storie i cui soggetti sono gli organismi viventi. La definizione e la decodificazione della sequenza dei quattro miliardi di nucleotidi che ci determinano è stato lo scopo del Progetto Genoma Umano. Leggere queste sequenze di lettere e decifrare i testi che esse compongono è oggi chimicamente possibile. Interpretarli è più difficile, richiede uno sforzo di tipo nuovo: il DNA è infatti un Testo senza Sommario. È organizzato in unità funzionali (i genomi) che hanno senso compiuto soltanto nella loro interezza, esattamente come i testi letterari.

Decifrazione del materiale genetico

La matrice chimica fatta di tante piccole unità disposte una accanto all'altra e sulla quale è scritto tutto quello che siamo ci è stata consegnata in due copie dai nostri genitori, una copia dalla madre, una copia dal padre. A nostra volta, forse, ne passeremo copia ai nostri figli.

Decifrazione del genoma umano significa definire la qualità e l'ordine delle piccole unità che compongono questa matrice. Il compito è complesso, perché negli esseri umani le unità sono più di quattro miliardi, e vanno definite una ad una. Per farlo è necessario svolgere una reazione chimica dopo l'altra, raccogliere i dati, ordinarli, confrontarli, studiarli, ripetendo queste operazioni milioni e milioni di volte.

Questa serie di operazioni porta alla definizione chimica del genoma, non a quella funzionale.

Nel *determinare la sequenza* viene definito l'ordine delle lettere chimiche che la compongono (di nuovo le A,T,C,G) ma non ne viene interpretato il senso. Una volta nota la sequenza delle lettere, è necessario capire come queste lettere sono raggruppate e quale è il significato dei messaggi che esse codificano. Il linguaggio di tipo semiologico-strutturalistico che stiamo usando qui è il linguaggio usato normalmente nell'affrontare questi problemi. È un linguaggio appropriato, poiché i problemi interpretativi posti dal DNA sono effettivamente simili a quelli presentati da una lingua sconosciuta, da un testo criptato, dalla interpretazione critica di un testo letterario.

Ci sono dunque due momenti: quello della genomica, la determinazione delle lettere, del loro ordine, della loro organizzazione in insiemi (i cromosomi) e in soprainsiemi (i genomi). E quello della post-genomica, il momento della interpretazione dei significati e delle loro multiple ed alternative integrazioni. Alla fine di questi due processi conosceremo con esattezza il nostro aspetto più intimo, non avremo più alcun segreto che ci nasconda a noi

stessi. Potremo capire “che cosa” realmente passiamo ai nostri figli, che cosa abbiamo ricevuto dai nostri genitori. Che cosa in termini di realtà genetica, non di metafora o di morale o di oggetti di famiglia. Potremo renderci conto che siamo combinazione ordinata di elementi che si tramanda come tale; non soltanto come materia ma, soprattutto, come *ordine* nell’organizzazione della materia.

Il concetto di “Testo” riferito al DNA

Gli organismi viventi sono organizzati intorno a testi genetici. La natura comunica attraverso testi. “Testo” indica “una catena di definizioni collegate da legami di coerenza”, non importa se complessi o semplici. Il miagolio di un gatto è un testo, la Divina Commedia è un testo. Un importante corollario di questa definizione semantica è che i testi esprimono significati diretti o indiretti, che sono normalmente definiti come significati lessicali e testuali, rispettivamente. Un oracolo greco, l’oracolo di Delfi, ad esempio, trasmetteva messaggi in forma di parole mescolate che *dovevano* essere interpretate, i testi che trasmetteva erano puramente *testuali*. Un segnale stradale, al contrario, è più diretto possibile, è puramente *lessicale*. La differenza tra i due tipi di testi non risiede nel loro grado di ambiguità ma nell’uso che ci si aspetta ne venga fatto.

Nel DNA le definizioni sono chiare e dirette, sono tutte lessicali; quando è preso nella sua totalità, però, il messaggio diventa testuale perché per sua natura si adatta al suo ambiente e risponde agli stimoli esterni (che per definizione sono enormemente complessi, data la loro infinita serie di possibilità alternative). Il DNA è fatto per adattarsi (ed anche, più precisamente, per adattare gli organismi che codifica) a tutte le situazioni possibili, a tutte le varianti concepibili. In questo senso funziona perfettamente, visto che è in grado di codificare organismi adatti a tutti gli ambienti del globo, presenti, passati e (spe-

riamo) futuri. Il DNA, come testo, può essere scritto in tutte le lingue concepibili. Il DNA come linguaggio è in grado di esprimere tutto, è in grado di “lottare con l’inesprimibile finché non arriva ad esprimerlo” (Kirkegaard, a proposito di altro, ovviamente). Una proprietà fondamentale del DNA, accanto a quelle della sua capacità informazionale e a quella della sua organizzazione, è la sua adattabilità ad un ambiente che cambia: la sua capacità di evolvere.

La distinzione tra significato diretto (lessicale) e indiretto (testuale), è uno dei problemi aperti della semantica, e riguarda da vicino il DNA. Per capire il significato di un testo, l’interprete non può non ricorrere a metodi di interpretazione integrata perché (i) il linguaggio del messaggio (ogni linguaggio ed ogni messaggio) è degenerato (molti sensi per una parola singola) e (ii) perché il messaggio non viene mai da solo. Quindi il punto è: quale è il senso della distinzione tra lessicale e testuale, quando ogni parola ha un significato lessicale e una interpretazione testuale? (Per la trattazione originaria di questi concetti, v. U. Eco, “Trattato di semiologia”). La soluzione è quasi ovvia: il significato lessicale è convenzionale e *diventa* situazionale nel processo di interpretazione da parte del destinatario. Nei sistemi biologici i meccanismi sono gli stessi: un ormone ha un ruolo convenzionale che, nella complessità dell’organismo, diventa situazionale.

Così il DNA. Tutti i suoi componenti, i suoi geni, hanno significati specifici, ben definiti, e questi cambiano a secondo dell’uso che se ne fa. Nelle nostre cellule, nei nostri cromosomi, c’è un numero altissimo di componenti genetici, come nelle lingue che parliamo c’è un enorme numero di *lessemi*, di unità di lessico.

La vita è l’uso che facciamo, che siamo costretti a fare, di questi geni. La lingua che parliamo è l’uso che facciamo, che siamo capaci di fare, dei lessemi che abbiamo in mente. Se vogliamo paragonare la lingua dei nostri geni più alla lingua scritta che a quella parlata (paragone forse

più coerente, visto che i geni sono codificati, sono “scritti”, nel DNA, e visto che il DNA è un filo che si conserva, è un nastro, un volume, non un discorso articolato da suoni che si disperdono), se vogliamo paragonare i nostri geni ad un testo scritto, è lecito domandarsi: di che natura è questo testo? È una commedia o una tragedia, un racconto o un dialogo? *Il DNA è Opera Combinatoria, opera illimitata, espressione della capacità di autoorganizzazione della materia: autogena, intrinseca, autoproduttore, illimitata.*

Riassumendo: secondo un approccio semiologico, i geni sono strutture lessicali, nel loro insieme assumono una valenza testuale. Questo tipo di confronto a prima vista arbitrario e in qualche modo forzato, è al contrario utilissimo per capire il significato vero della struttura organizzativa dei genomi, la loro evoluzione, il loro meccanismo di funzionamento di base.

Ragionare sulle opportunità, sulla liceità e sulle conseguenze degli interventi sui genomi richiede di conoscere le regole organizzate delle strutture che vogliamo modificare. Più che in termini di *germinale* e *somatico*, la nostra struttura genetica è organizzata in insiemi *lessicali* e *testuali*. È importante, se spostiamo un gene, sapere se stiamo cambiando solo la valenza lessicale e se al contrario stiamo alterando un valore testuale.

Alterare un testo è operazione di maggior peso che cambiare il modo di usare una parola.

Alterare un gene è operazione di maggior peso che cambiare una breve sequenza di nucleotidi.

L'unità del Testo

Nel tentativo di capire la natura della organizzazione di

un genoma non dobbiamo limitarci a considerare dove le frasi genetiche (i cromosomi) iniziano e finiscono, non dobbiamo rimanere nel contesto di un sistema interpretativo costruito intorno a comparazioni ingenuie e semplici in cui i geni che codificano le proteine sono i soggetti o i nomi delle frasi, in cui le strutture che regolano l'attività genetica sono i verbi (perché indicano azione), in cui le strutture che ne stimolano la funzione sono gli aggettivi (perché regolano l'intensità, la quantità, le proprietà temporali e quantitative dei geni-nomi), in cui i prodotti dei geni che agiscono su altri geni sono avverbi e congiunzioni, usati per correlare le varie parti della storia dell'individuo. Questo tipo di comparazioni è certamente utile, permette lo sviluppo di paralleli e di similitudini, stimola la messa a fuoco e lo sviluppo di nuove idee ma soffre, inevitabilmente, di frammentazione.

Dobbiamo invece guardare al DNA come ad un oggetto unico, nella sua interezza, nel suo equilibrio, DNA come un testo che acquista il proprio significato solo quando viene letto da un'estremità all'altra.

Esistono esperimenti molto interessanti fatti in *Salmonella* nei quali una parte del grande cromosoma circolare di questo batterio veniva spostato da un punto all'altro della sua mappa genetica. I geni rimanevano gli stessi ma il loro ordine reciproco era stato cambiato. Apparentemente le *Salmonelle* non soffrivano per il rimescolamento subito, crescevano con la stessa efficienza di prima della manipolazione.

Dopo qualche generazione, però, si osservava che i geni erano tornati spontaneamente ai siti originali, come se conoscessero il proprio punto di origine, come se conoscessero l'ordine preferenziale che forma lo schema della loro vita normale. Quest'ordine non ha per noi senso, non troviamo alcuna correlazione apparente con l'espressione temporale dei geni, né con il flusso delle funzioni metaboliche. Ma la realtà sperimentale mostra che un ordine esiste. Il fatto che noi non comprendiamo la

logica dell'ordine usato all'interno del testo del DNA mostra che non siamo in grado di pensare come pensa una cellula; o almeno non ancora.

L'avvenuta decifrazione

Il lunedì 26 giugno 2000 rappresenta una pietra miliare per la biologia e la medicina. L'annuncio fatto della avvenuta decifrazione del genoma umano stabilisce senza dubbio un punto di arrivo e allo stesso tempo un punto di partenza formidabili. L'interesse suscitato è stato giustamente enorme, non sono stati posti quesiti etici, quello che è stato riportato è un fatto di pura conoscenza. I frutti possibili di questa conoscenza non sono stati ancora raccolti, né sono stati valutati pienamente, né sono stati compiutamente messi a fuoco i problemi che questa conoscenza apre.

In primo luogo la decodificazione non è, e non sarà mai, veramente completa. La raccolta dei dati e la loro organizzazione è avvenuta da parte di due strutture distinte ed ha seguito due approcci non identici. L'organizzazione pubblica (HUGO, Human Genome Organization) ha visto impegnati laboratori di molti paesi differenti, singolarmente o in gruppi di collaborazione. L'organizzazione privata (Celera Genomics di Rockville, Maryland) ha invece lavorato da sola, essenzialmente usando scorciatoie e utilizzando *anche* i dati che l'organizzazione pubblica metteva man mano a disposizione. In linea generale il risultato della HUGO è più solido e più completo, quello della Celera è più snello, più rapido, e dal punto di vista della qualità dell'informazione ottenuta, è altrettanto buono.

Una differenza sostanziale è che i dati della HUGO sono pubblici ed accessibili, quelli della Celera sono ottenibili ma hanno un prezzo. Queste due affermazioni non hanno però, così come formulato, un significato comple-

to e non sono riducibili ad una situazione semplice. Vanno valutate alla luce di *che cosa* è effettivamente disponibile e di *come* le informazioni sono state ottenute.

Prima di tutto la *validazione*. Il processo di determinazione delle sequenze ha in sé possibilità di errore. Altri errori vengono introdotti dal procedimento di elaborazione dei dati. La sicurezza della veridicità del prodotto finale è ottenuta attraverso la ripetizione del procedimento. La ripetizione da parte della HUGO è di sette volte, quella di Celera è di cinque (in media). Celera inoltre ha potuto verificare i dati contro quelli pubblicati da HUGO, mentre il contrario non è stato possibile.

La differenza sostanziale è però dovuta alla natura stessa della organizzazione dei genomi: quello che ha maggior interesse in un genoma sono i suoi geni, definiti come i tratti di DNA che contengono l'informazione (codificano) determinati prodotti. Questi prodotti sono soprattutto proteine o RNA funzionali (in genere RNA usati per il trasporto di aminoacidi ed RNA che compongono delle strutture – i ribosomi – fondamentali per la sintesi delle proteine stesse). I genomi di organismi semplici sono costituiti quasi esclusivamente di sequenze che codificano geni; altre sequenze (quelle di tipo organizzativo-strutturale, quelle di tipo regolativo, quelle che determinano i punti di inizio della replicazione, e così via) sono presenti in modo quantitativamente minoritario e spesso sono sovrapposte ad altre funzioni. I tratti che determinano strutture e regolazioni sono nei virus, nei batteri e nei funghi inferiori, ridotti al minimo.

Negli organismi superiori, al contrario, la quantità di DNA impegnata in funzioni geniche è minima; nell'uomo non supera il 3%. Il resto ha molteplici funzioni e forse in parte non serve a nulla (il concetto di "*servire a*" è naturalmente molto antropocentrico e andrà analizzato a parte, insieme al concetto di "*gene egoista*"). Comunque, il DNA direttamente codificante è poco, brevi tratti immersi in un gomitolo enorme di sequenze.

A questo punto è possibile capire in che cosa differiscano i due approcci. La HUGO ha determinato *tutta* la sequenza (o quasi): partendo da un punto ha cominciato a leggere, indipendentemente da cosa incontrava. Il processo è laborioso e fornisce informazioni rilevanti solo *dopo* che la sequenza è stata letta, organizzata, analizzata per contenuto informativo. A quel punto l'informazione sarà di qualità alta perché non solo sapremo come è fatto quel determinato gene, ma anche come è fatto il tratto di cromosoma nel quale quel gene è contenuto. Naturalmente, poiché siamo partiti da una situazione di grande ignoranza della mappa genetica ad alta risoluzione, i risultati sono stati in buona parte inattesi e di grande rilevanza interpretativa. Il problema è che spesso non è possibile attribuire funzioni a sequenze che non hanno fenotipo identificabile, sequenze cioè la cui mancanza o alterazione non comporta modificazioni apprezzabili nell'organismo considerato, oppure sequenze che non hanno un corrispondente identificabile in un altro organismo evolutivamente simile. Sequenze che *sembrano* non servire a nulla.

L'approccio della Celera è speculare: partendo dai prodotti genici ha fatto la strada a ritroso. Il procedimento è più rapido, più facilmente utilizzabile, enormemente più incompleto, finalizzato all'uso (e non alla comprensione) della struttura. L'incompletezza dei dati ottenuti è stata risolta in parte ricorrendo allo sviluppo massiccio di tecniche di bioinformatica e di automatizzazione della procedura di sequenziamento. Il procedimento va sotto il nome di "shotgun", sparo nel mucchio. Va al merito di Craig Venter aver sviluppato con grande credibilità un metodo che veniva giudicato inadatto per la sua supposta imprecisione intrinseca. Venter ha puntato sulla capacità di sviluppo autocatalitico della bioinformatica, ed ha vinto. Una volta identificato un tratto contenente informazione di codificazione (un tratto cioè *codogenico*), il computer accoppia a quella sequenza i tratti di DNA

sequenziati “a caso”, valutando statisticamente la validità del posizionamento ottenuto. Semplificando un po', venivano aggiunte tessere di puzzle estratte a caso intorno a pochi punti di partenza noti e posti all'interno dello schema. La HUGO invece ha costruito prima i bordi del puzzle ed ha poi proseguito con ordine.

In entrambi i casi i risultati sono di ottima inattesa qualità, sono ancora incompleti (e questo giustifica le affermazioni tipo: “abbiamo *finalmente* finito”, che vengono ripetute di tanto in tanto). E lo saranno a lungo perché la ripulitura delle zone complesse e la definizione completa dei tratti ridondanti è non solo tecnicamente molto difficile ma è considerata meno interessante. Una conseguenza delle ambiguità e della mancanza di attribuzione di significato è l'importanza dello studio di genomi di riferimento e di comparazione (leggi “genoma del topo”) sui quali siano possibili interventi genetici sperimentali.

***Finestra 3* Analisi di espressione genica a livello globale**

Tecnica che permette di determinare, all'interno di tutti i geni presenti nel genoma in considerazione, tipicamente in uomo o topo, quale gene venga espresso in una data situazione differenziativa. La tecnica, nota con il nome di "Microarrays display", rende possibile sapere quali geni sono espressi, tra tutti i geni del corredo genomico, in una cellula specifica, in una condizione genetica o fisiologica definita. In alcuni organismi è possibile già oggi conoscere le risposte per tutti i geni presenti, in altri l'analisi si può fare solo per geni selezionati e comunque rappresentativi di tutte le classi geniche principali.

3. Dall'informazione all'uso

Il significato vero della decifrazione del genoma è l'aver reso possibile la definizione del programma di assemblaggio e mantenimento di un essere umano, programma presente in ogni momento della vita in ogni cellula dell'individuo (meno che nei suoi anucleati globuli rossi). Accanto alla prospettiva generale è resa anche possibile l'analisi specifica, locale, di come funzionino (o mal-funzionino) le variazioni allo schema generale.

Il cambiamento di un singolo nucleotide all'interno di un gene fatto di decine di migliaia di queste unità può costituire una mutazione che comporta una alterazione rilevante di una data proteina e, se questa proteina è importante, l'organismo smette di funzionare. Oppure il cambiamento di quel nucleotide non ha alcuna importanza, cade semplicemente nell'ambito delle tante alternative possibili (un poliformismo, un SNP, vedi il paragrafo specifico più avanti) o può, perfino, comportare un miglioramento.

Una volta identificati i geni, il compito diventa quello di identificare tutte le loro funzioni, le funzioni dei loro cambiamenti ed i loro legami con le malattie.

Funzioni semplici e funzioni complesse

In molti casi è valida l'equazione: un gene = una funzione. E per verificare l'equazione se ne cerca la variazione e/o l'alterazione, si cerca il caso in cui si ha: un gene (alterato) = una malattia. Ricercatori impegnati nell'Human Genome Project hanno identificato decine di geni coinvolti in sindromi precise, spesso scoprendo o confermandone il rapporto diretto e univoco. Un tipo di emofilia, ad esempio, è causata da un unico gene difettoso ed è stato facile identificarlo attraverso studi di trasmissione familiare. In molti altri casi – i princi-

pali: malattie cardiache e cancro – i geni coinvolti sono molti ed il rapporto tra loro è certamente molto complesso; tenendo anche presente che la loro funzione è nei processi patologici mascherata, o modulata, o esaltata dalle interazioni con l'ambiente. Alla decifrazione della sequenza ed alla mappatura dei geni si sovrappone quindi la necessità di studi di popolazione. La validità di questo approccio crociato è stata dimostrata dalla iniziativa ambiziosa della deCODE.

A capitale americano ma con sostanziali interessi in loco, la deCODE ha ricostruito le mappe genetiche di potenzialmente *tutta* la popolazione islandese. Da un lato si ritiene che i 270.000 individui che compongono la popolazione stabile siano di origine genetica sufficientemente omogenea, dall'altro la stabilità della popolazione in pochi centri abitati e la natura stessa della loro struttura sociale ha permesso il mantenimento della documentazione e la possibilità della ricostruzione molto indietro nel tempo dei rapporti di parentela. Su questa enorme mappa si inseriscono i dati di sequenza dei geni che vengono studiati e si identificano così le varianti patologiche. Lo studio ha permesso l'identificazione di geni collegati alla schizofrenia, al diabete ed a molte affezioni del sistema cardiovascolare. La costruzione di questi dati è molto laboriosa: i risultati veri vengono raggiunti dopo un periodo abbastanza lungo. Gli islandesi hanno aderito in larga misura all'iniziativa: il beneficio che ottengono in cambio della loro storia genetica passata e delle informazioni sulla sequenza del DNA presente che vive in loro oggi non è economico. Hanno ed avranno in cambio diagnosi e prodotti terapeutici specifici.

Il potenziale beneficio in termini di medicina collettiva (ed in termini economici) è apparso subito chiaro. Il progetto impostato su scala così ampia in Islanda è stato imitato in due regioni canadesi, Labrador e Terranova. In questo caso i termini sono diversi: la popolazione è molto più dispersa ed eterogenea. La mescolanza di geni di pro-

venienza tanto disparata (eschimesi, francesi, anglosassoni, amerindi) e la base filologica di partenza più limitata (a volte si risale all'indietro per non più di due o tre generazioni, spesso per nessuna) rende il confronto tra i due esperimenti molto interessante. In qualche modo l'esperimento canadese è rivolto più alla costruzione di banche dati future che all'interpretazione dell'eredità che ci viene dal passato.

L'iniziativa della deCODE non è però priva di aspetti controversi. Il data-base dovrebbe essere completamente anonimo per poter garantire la riservatezza non solo riguardo i possibili problemi genetici di un determinato individuo o di una determinata linea familiare; la riservatezza è soprattutto importante per garantire il presente (malattie in atto non ancora diagnosticate, possibilità non ancora avverate ma definibili in termini di probabilità) e il futuro tendenziale, ad esempio, di criminalità o intelligenza. Questo tipo di considerazioni è stato alla base della contestazione dell'accordo deCODE-governo islandese da parte della Mannvernd, una associazione di medici e cittadini basata a Reykjavik. Il punto sollevato è di incostituzionalità perché, si è affermato, contraddice l'assunto di base del rapporto medico-paziente. Da un lato quindi c'è il problema che i dati del Genoma Umano sono potenzialmente fondamentali ma oggettivamente poco utili se non calati in una realtà specifica fatta di cartelle cliniche e di identificazioni genetiche individuali; dall'altra la necessità di garantire una difficile riservatezza: il nome di chi si ammala o muore va annotato in una scheda e, per definizione, l'individuo non può essere anonimo. L'insieme di problemi non è reso più semplice dal fatto che i dati vengono raccolti per silenzio-assenso: chi non vuole entrare nella rete del data-base deve dichiararlo. Né dall'accordo commerciale tra la deCODE e la Compagnia farmaceutica svizzera Hoffmann-La Roche che ha acquistato per 200 milioni di dollari il diritto di consultazione del data-base.

Benché su scala molto minore, è anche interessante l'esperimento in corso sulla popolazione dell'isola di Norfolk, al largo dell'Australia. Norfolk è l'isola sulla quale si rifugiarono i marinai che si erano impadroniti del Bounty nel 1789. Dopo aver rilasciato in mare il capitano William Bligh, gli ammutinati e le loro donne polinesiane hanno cominciato a riprodursi. Oggi i 990 abitanti dell'isola sono, dopo più di 200 anni di isolamento, sotto esame genetico. Il profilo di interazioni genetiche (ufficiali) è completamente ricostruibile e la dieta è stata ed è molto costante. Sia i polinesiani che gli anglosassoni soffrono di malattie cardiocircolatorie. I tipi specifici di sindromi sono però in parte diversi. L'analisi di quali geni siano coinvolti in un tipo o nell'altro di patologie, i pedigree genetici intrecciati, i dati sulle età e cause dei decessi, rendono il laboratorio di Norfolk di estremo interesse. I primi dati sul coinvolgimento di geni determinati con la predisposizione all'emigrania ed alla alta pressione sanguigna sono stati già raccolti. È facile prevedere che questi esempi di meta-genomica di popolazione saranno presto seguiti da altri. Il caso di quanto è in atto in Inghilterra è particolarmente interessante. E sarà anche interessante vedere cosa i gruppi sociali che cedono le proprie preziose ed irripetibili informazioni genetico-storiche avranno in cambio.

Genomica comparata

Benché si parli oggi più di epigenomica che di genomica, per poter capire cosa si intenda con la prima, dobbiamo prima sapere in cosa consista la seconda. La scoperta che soltanto una piccola percentuale (tra il 2 ed il 3 per cento) del genoma umano codifica geni è giunta insieme a quella che il loro numero non è superiore a 25.000. Abbiamo cioè molti meno geni di quelli che il nostro orgoglio umano si sarebbe atteso: il numero è sostanzialmente

uguale a quelli del topo (altro mammifero per il quale abbiamo stime molto precise), e poco più del doppio dei geni della *Drosophila*, il moscerino dell'uva. La sostanziale identità genetica tra uomo e topo apre due prospettive molto importanti: da un lato la organizzazione di un archivio di topi mutanti, ognuno diverso per uno o più geni dal tipo selvatico normale, dall'altro la verifica sperimentale degli effetti di queste mutazioni (ottenute soprattutto con tecniche di transgenesi) e delle terapie corrispondenti. Archivi amplissimi di topi transgenici sono stati organizzati, spesso come risultato di cooperazione sovranazionale, ed i risultati in termini di identificazione gene-patologia sono stati immediati. È risultato inoltre assolutamente irrinunciabile lo sforzo di decifrare prima possibile l'intero genoma del topo. Questo sia per permettere l'identificazione del gene mutato con una funzione definita, sia per identificare quale parte del gene nel topo (e quindi del suo corrispondente umano) è coinvolta nella malattia. Il presupposto della validità e necessità di questo approccio è che tra uomo e topo non solo sono conservati il numero ed il tipo specifico dei vari geni, ma ne è sostanzialmente conservata anche la qualità di sequenza.

La funzione della proteina che svolge un certo ruolo è ottenuta con una determinata struttura. Poiché la struttura della proteina è evolutivamente molto conservativa, sarà mantenuta anche la sequenza informativa del gene specifico.

Il risultato principale di questi studi è che in generale le strutture (ed i geni corrispondenti) sono molto antiche, molto più antiche sia dell'uomo che del topo e sono state inventate in organismi primigeni che non esistono più. *Geni e strutture sopravvivono in noi, permettendoci di esistere e svolgere le nostre poco originali funzioni fisiologiche e riproduttive.*

L'importanza di conoscere il genoma del topo altrettanto bene di quello umano è risultata chiara sia all'am-

biente scientifico internazionale che alla Celera Genomics e al suo fondatore/padrone Craig Venter. Celera ha sequenziato più del 95 per cento del genoma murino; naturalmente con il suo sperimentato approccio metodologico che gli ha consentito, tra l'altro, di chiederne il brevetto d'uso man mano che la identificazione delle sequenze procedeva; man mano che procedeva, come si dice, il loro "tagging".

Le analisi genomiche *in toto* riguardano ormai moltissimi altri organismi. Dopo i sequenziamenti, datati ora di più di un decennio, di nematodi, dello Zebrafish, del pufferfish (*Fugu rubripes*), di lieviti, del riso e di molti batteri di interesse medico o biotecnologico, sono stati portati a termine sequenziamenti di migliaia di specie, da quelle a rischio estinzione, a quelle filogeneticamente nodali, a quelle patogene. È anche stata in particolare completata la sequenza del verme *Caenorhabditis elegans*. Può sfuggire ai più l'interesse di investire energie e risorse per sequenziare un verme. La sua rilevanza è nel fatto che *Caenorhabditis* ha un numero di cellule basso e definito, e che si conosce il fato e la funzione di ogni sua cellula individuale, singola e in relazione con le altre. Questo ne fa un sistema modello ideale per studiare fenomeni fondamentali quali l'apoptosi (il destino terminale autoimposto delle cellule), l'imprinting differenziativo a livello molecolare (perché, quando e come una cellula sa che deve diventare differente dalle altre), il ruolo delle cellule staminali (analizzabili qui con particolare facilità perché di pochi tipi), i meccanismi molecolari dell'apprendimento (ovvero, in termini fisico-chimici: ricordare, reagire). Ce n'è abbastanza per giustificare gli sforzi compiuti per guardare in dettaglio la doppia elica di questo umile piccolo verme, breve in confronto a quella di altri organismi, sulla quale tutto questo avviene.

Una volta organizzata la struttura e addestrato il personale, impostare studi di sequenziamento è relativamente agevole e potenzialmente molto remunerativo.

Considerando la ricaduta farmacologica, la possibilità di brevettare interi organismi in cui sia stato inserito un gene nuovo purché la sua sequenza sia nota, o semplicemente l'ammortamento delle spese iniziali, il futuro prossimo biotecnologico risulta indissolubilmente legato a studi genomici.

Un caso illuminante di questo punto è la grande messe di dati che sono stati ottenuti per i genomi di patogeni classici: *Microbacterium leprae* (il batterio della lebbra) e *Microbacterium tuberculosis* (che causa la TB), ad esempio. Il confronto dei due genomi ha fornito dati importantissimi sullo sviluppo di metodi di diagnosi precoce e semplice, e di cure efficienti e risolutive.

Altri Genomi

Drosophila melanogaster è il moscerino della frutta, sistema modello per definizione della scienza genetica. Le ragioni che lo hanno reso tale sono le stesse che hanno fatto del lievito *Saccharomyces cerevisiae* il modello di studio della genetica e della biologia molecolare indirizzate alla produzione biotecnologica. Le drosofile svolazzano nelle Università di tutto il mondo da più di un secolo. La loro genetica è nota in grande dettaglio, il loro genoma è stato sequenziato. L'annuncio del quasi completamento del suo sequenziamento è venuto a metà febbraio 2000, da parte di Craig Venter che parlava questa volta a nome di un consorzio che raggruppa il progetto internazionale centrato a Berkley: *Drosophila Genome Project* (BDGP) ed il corrispondente progetto europeo (EDGP). Sono stati sequenziati 13.601 geni (il 97% del totale), 120 milioni di basi (la maggior parte del contenuto di DNA di questo genoma), si è scoperto che 40 milioni di queste basi sono "junk DNA" (DNA residuale senza apparente funzione), che il 60% dei geni umani noti essere correlati a situazioni patologiche hanno un corrispon-

dente in drosofila. Particolarmente rilevante sarà studiare, usando i raffinatissimi metodi genetici sperimentali disponibili in questo sistema, i rapporti tra i numerosi geni coinvolti nel cancro e la integrazione delle loro varie funzioni.

Un fatto molto interessante è l'aver scoperto i geni specifici della drosofila, il loro gran numero. Cosa significa essere drosofila è scritto in queste sequenze che nessun altro organismo possiede, che i moscerini hanno inventato per costruirsi antenne che nessun altro ha, e riconoscere odori che nessun altro sente.

Ma se guardiamo nell'intimità della doppia elica del topo, la situazione è differente. Il suo DNA è talmente simile al nostro che capire in termini genetici cosa significa essere topo o essere uomo richiede uno sforzo di immaginazione notevole. Il progetto di sequenziamento di *Rattus norvegicus* è stato intrapreso subito dopo quello del sequenziamento del genoma umano. Si è anche scoperto che, nonostante le apparenti differenze, *Mus* e *Homo* hanno genomi molto simili: in entrambi un numero simile di nucleotidi, meno del 3% dei quali sono geni, il resto sequenze residuali. Le sequenze geniche sono inoltre molto conservate. Considerando il valore aggiunto offerto dalla possibilità della sperimentazione genetica sul topo, la conoscenza del genoma di *Mus* è diventata l'indispensabile Pietra di Rosetta della ricerca post-genomica.

Single Nucleotide Polymorphisms (SNPs)

Non siamo tutti uguali. Il "genoma sequenziato" non è quello di una persona specifica, ma viene da un insieme di donatori scelti e mescolati in modo molto mirato: da 28 donatori di DNA per HUGO, da 5 per Celera. Quando gli è stato chiesto, Venter ha rivelato le sue fonti: 2 caucasici, un ispano-americano, un negro, un indiano; alla domanda se uno dei caucasici fosse lui non ha commentato.

Quando questo Genoma di riferimento viene confrontato con quello di un individuo definito, si trovano immediatamente differenze, come ovvio e atteso se guardiamo un piccolo numero di persone radunatesi a caso: non siamo tutti uguali, né fenotipicamente né genotipicamente. La differenza a livello genomico tra due individui presi a caso è solo dello 0,1 per cento. Questo significa che, dati 4 miliardi di nucleotidi di partenza, due umani sono diversi in 1 milione di punti singoli. Questo carattere, detto SNP (Single Nucleotide Polymorphism) è molto rilevante e, di nuovo, è stato oggetto di una corsa di competizioni tra ricerca pubblica-privata e settori esclusivamente privati. Alla fine del 2000, dopo un rush iniziale, erano stati mappati (e resi pubblici) da parte di un consorzio formato da cinque Università, dalla Wellcome Trust e da 10 case farmaceutiche, 800.000 siti SNPs. La Celera ha contemporaneamente posto in vendita l'accesso ad una banca dati comprendente 2.8 milioni di SNPs. Da allora il loro numero è in lenta, costante crescita.

Gli interessi principali vertono sui tre aspetti seguenti: analitico-predittivo, analitico-farmaceutico, drug-tailoring. Il primo è di facile comprensione: ammettiamo di avere a disposizione un microchip che porti sulla sua superficie una serie di punti analitici, disposti secondo un ordine noto ed analizzabili in modo automatico. I punti analitici corrisponderanno ognuno ad un singolo diverso polimorfismo SNP relativo ad un gruppo di geni rappresentativo di caratteri di predisposizione ad un gruppo di malattie o a "tendenze" rilevanti per le caratteristiche comportamentali di un individuo. In molti casi il rapporto tra un certo gene e le sue varianti è stato definito in modo preciso; ad ogni SNP di un gene che sia coinvolto, diciamo, nella schizofrenia corrisponderà un segnale elettronico preciso. Mettiamo sul chip un campione di DNA dell'individuo da analizzare, e il gioco è fatto. In tempo brevissimo avremo la risposta, e la precisissima caratterizzazione genetica delle persone. Questi

“DNA gene-chips” sono disponibili; quelli prodotti dalla Affymetrix, ad esempio, sono per uso umano e possono leggere migliaia di varianti contemporaneamente. Altre società producono DNA gene-chips per analizzare contemporaneamente tutti i geni dei principali organismi patogeni noti, altri chips sono specifici per le malattie che riguardano il sistema immunitario, e così via. In presenza di tante determinate variabili (il cui insieme è la nostra identità genetica) e proprio a causa della loro automatizzata determinazione, classificazione, confronto, è facile immaginare come il problema della privacy genetica stia diventando sempre più centrale. È anche importante aver chiaro che questi chips rappresentano *la metodologia con la quale capire il meccanismo di lettura dell'Enciclopedia che è il nostro genoma. Se non possiamo leggerlo, aver trascritto il libro del nostro DNA è inutile.*

Il livello analitico-farmaceutico ha trovato nello studio degli SNPs la risposta al relevantissimo problema della differenza di reazione ai farmaci. Alcune sostanze terapeutiche funzionano in determinati individui e non in altri, secondo uno spettro di possibilità e di variazioni molto complesso, al punto che negli Stati Uniti questo problema è stato stimato essere la quarta causa di decesso. Esempi chiarissimi sono descritti nel campo degli antibiotici, degli antitumorali, dei β -bloccanti, degli steroidi. La soluzione è offerta dalla determinazione del rapporto tra profilo SNPs ed attività del farmaco. I progressi in questo campo sono talmente rapidi e talmente rilevanti che nel dibattito che li accompagna le obiezioni basate sulla difesa della privacy vengono a mala pena prese in considerazione.

Nella stessa direzione va il rapporto tra SNPs e programmazione di farmaci. Quale è il cambiamento strutturale nella proteina prodotta da un gene alterato in una data situazione SNPs? E quale sarà il farmaco che riporta questa proteina alla giusta struttura? Ci sono pochi

dubbi che in un futuro prossimo (data una situazione di efficienza e disponibilità di risorse) la casa farmaceutica produrrà il farmaco, mentre l'individuo dovrà mettere a disposizione il proprio profilo SNPs per ottimizzarlo. È anche sotto questo aspetto che va inquadrata l'iniziativa della deCODE che, nel suo processo di costruzione del-

***Finestra 4* Tecniche di base: la costruzione di DNA gene-chips ed il Microarrays display**

I primi sono in grado di analizzare in pochi minuti ed in modo completamente automatico le varianti genetiche (eventualmente patologiche) dell'individuo in esame oppure sono in grado di identificare che tipo di agente patogeno è presente nel campione esaminato. L'automatizzazione elettronica di rivelazione, la produzione robotizzata dei DNA gene-chips e la semplificazione del processo di ibridazione renderanno, nel prossimo futuro, il costo della strumentazione accessibile anche a livello periferico. La versatilità dei DNA gene-chips è altissima: si possono inserire ed esaminare contemporaneamente grandi quantità di sequenze. La scelta di cosa analizzare è favorita da una serie di successi tecnico-organizzativi acquisiti di recente: la determinazione completa (o quasi) del genoma umano, la creazione delle banche dati corrispondenti, la compilazione di cataloghi SNPs, l'associazione tra caratteri genetici e patologie definite. E' anche chiaro l'aspetto commerciale del sistema: le informazioni hanno un prezzo, l'aggiornamento dei cataloghi è una necessità continua, opera aperta. L'altra tecnologia, i Microarrays, discende dalla prima. Invece di esaminare direttamente i geni questa tecnica analizza il profilo della loro espressione su tutto il genoma. Se consideriamo che soltanto pochi geni vengono espressi in un dato tessuto in un dato momento, comprendiamo subito che una alterazione del profilo di espressione genica rivela immediatamente un eventuale stato patologico, con grande dettaglio molecolare.

l'albero genealogico della intera popolazione islandese, ha mappato sia mutazioni che SNPs.

4. *Faber fortunae suae*: Genetica ed Epigenetica

In termini genetici, ognuno di noi è *faber fortunae suae*: viviamo determinati da una equazione i cui termini sono *libertà e responsabilità*. Noi e il nostro DNA, dunque, noi e il DNA che abbiamo ereditato, con il quale partiamo per il nostro viaggio al momento del concepimento e che, alla fine, restituiamo. Detto con Cicerone: *Ut navem, ut aedificium idem destruit facillime qui construxit, sic nomen eadem optume quae conglutinavit natura dissolvit* (De senectute 72), come una nave o un edificio li demolisce più agevolmente di ogni altro chi li ha costruiti, così l'uomo si dissolve più facilmente nelle mani della natura che lo ha composto. Sembra quasi una benpensante ovvietà, una benevola ripetizione dell'atteso. In realtà significa che chi vive in armonia con la propria natura sa di non dovere nulla ad altri se non ai propri avi, sa di non poter delegare se non a se stesso. Ritengo molto fortunata la mia generazione che ha potuto per la prima volta sapere in cosa consiste essere autogeniti, che può leggere la doppia elica del proprio DNA. Il DNA è una delle chiavi per aprire la porta della conoscenza di se stessi.

Faber fortunae suae: ognuno è responsabile in quanto individuo libero soprattutto se è, allo stesso tempo, in armonia con il proprio ruolo genetico e la propria posizione nello spazio-tempo della Specie alla quale appartiene: *senex, nec vero dubitat agricola, quamvis sit quarenti cui serat, respondere: "Dis immortalibus..."*, e il contadino, per vecchio ch'egli sia, se gli chiedi per chi semini non esita a rispondere: *"Per gli Dèi immortali..."* (De senectute 25).

Perché cioè, senza citar troppo latino, ... gli Dèi non possono aver voluto solo che ricevessi tutto questo dagli avi,

ma anche che lo trasmettessi ai posteri. Trasmettere una eredità è (anche) generosità. Almeno generosità genetica. Ovvero: siamo responsabili del DNA che ci è stato affidato al momento del concepimento. Traducendo Cicerone in termini di genetica molecolare: gli Dèi hanno voluto che io ricevessi tutto questo, il mio DNA, e che io lo trasmettessi ai discendenti. E ad esso, alla mia natura, non mi devo ribellare. Nel frattempo lo coltivo, lo nutro, con la mia vita lo modifico, lo rendo adatto all'ambiente in cui vivo. Ma come? E perché?

Poichè l'Epigenetica è l'uso che facciamo del nostro DNA, la sua importanza è ovvia.

Il *perché* è chiaro. I miei discendenti vivranno presumibilmente in un ambiente simile a quello nel quale vive il mio genoma ed il corpo che lo contiene e lo nutre. Comunque, il mio DNA non conosce altri ambienti. Ovvero, meglio, conosce l'ambiente che gli è presente, conserva traccia di dove e come hanno vissuto i suoi antenati, conosce le loro doppie eliche precedenti, ma non conosce il proprio futuro, non sa dove vivrà. Per poter capire il *come* sono necessarie le informazioni che verranno date nel prossimo Capitolo. È naturale a questo punto tracciare una prima distinzione tra Genetica ed Epigenetica, indicandone le differenze principali.

La Genetica è l'eredità dei caratteri del vivente ed i meccanismi che la permettono. In prima approssimazione la genetica corrisponde in gran parte al DNA. Abbiamo in questo lungo discorso ripercorso le tappe della ricerca recente che ha portato alla decifrazione di interi, numerosissimi genomi, cosa che ha cambiato radicalmente la nostra visione della sfera biologica nel suo insieme. Tutti gli organismi ricevono il proprio DNA, la propria informazione, dai genitori. Lo ritrasmetteranno leggermente mutato ai propri discendenti se ne avranno.

Una mutazione è il cambiamento permanente (è importante sottolineare questo carattere) della

sequenza di lettere chimiche che, nel loro insieme, costituiscono il filo del DNA e dell'informazione. Non è vero però che nasciamo nudi con il nostro DNA nudo. È tutto molto più complicato.

Epigenetica. Oltre ad essere una sequenza di sostanze chimiche, di per sé stabile ed ereditabile in quanto tale, il DNA può anche essere durante la propria vita modificato in modo non permanente. Questi cambiamenti possono essere in parte ereditabili. Queste modificazioni sono dovute a piccole molecole che interagiscono con il DNA e che su di esso iscrivono il vissuto dell'organismo. “*Su*” di esso, “*epi*”-genetica, appunto.

Oltre ad essere il depositario dell'informazione a lungo termine, il DNA ha dunque una memoria a breve termine, conserva traccia della propria esperienza, e dell'esperienza del resto del corpo. Lo scopo di questo straordinario meccanismo di gestione dell'informazione è quello di modulare in modo flessibile la risposta comportamentale momento per momento.

Il DNA impara e scrive su di sé la propria esperienza, la propria cultura. La trasmissione della cultura di cui è portatore l'individuo che di quel DNA è fatto è la forma più alta di Epigenetica. Il DNA crea i meccanismi per leggere la propria Biblioteca.

5. L'infinito DNA

Il nostro infinito biologico è il DNA. Esistono genomi semplici e genomi complessi, ed è ormai chiaro che il concetto stesso di genoma è in qualche modo arbitrario e forzato.

Il genoma minimo, si considera, è quello di alcuni virus, composto di tre geni soltanto: (i) un gene per la proteina di rivestimento, per la protezione cioè dell'acido nucleico che, a questi livelli di semplicità, è in genere un RNA; (ii)

un gene per una polimerasi specifica, per l'enzima cioè che dovrà replicare quell'acido nucleico, quel genoma; e (iii) un gene per una proteina di riconoscimento dell'ospite, per l'interazione selettiva cioè con quell'organismo che lascerà usare il proprio metabolismo dal virus parassita.

Questa richiesta minima corrisponde alla definizione di identità ed alla funzionalità necessaria per la sua oggettivazione. In termini di nucleotidi, questo livello di complessità corrisponde a 3000-5000 unità. All'altro estremo della scala troviamo i genomi vegetali, la cui complessità raggiunge i 100 miliardi (!!!) di nucleotidi. I principi organizzativi sono essenzialmente gli stessi, la biochimica sistemica è identica, ma la distanza tra qualche migliaio e 100 miliardi implica un uso dell'informazione sostanzialmente diverso. Ed implica anche la mancanza di un limite superiore alla complessità raggiungibile dalla organizzazione dell'informazione genetica.

Genoma significa dunque null'altro che *Insieme*. È un'organizzazione, una coordinazione, una selezione di funzioni scelte, tra le tantissime potenzialmente usabili, perché collaborino per il bene dell'organismo che determinano, tra le tantissime potenzialmente usabili. *Una Biblioteca che si organizza è una vita. Non è mai, diciamo, la somma di libri isolati* (Carlos M. Dominguez, La casa di carta).

Mancanza di limite superiore comporta mancanza di definizione, ed eccoci qui di nuovo a chiedere aiuto agli antichi e alla loro intuizione del concetto di *apeiron*. Che vuol dire indefinito e al tempo stesso indefinibile, e non vuol dire soltanto infinito. Infinito corrisponde a non misurato e non misurabile, mentre per *apeiron* manca sia la misura che la possibilità di porsi il problema. Il DNA è *apeiron*, perché la complessità potenziale non ha limiti: moltiplichiamo i 100 miliardi di nucleotidi di una cellula di *Lilium* per tutte le possibilità alternative di combinazione di sequenza (elevando cioè alla quarta), poi per

tutte le possibilità di ricombinazione, poi per tutte le varianti epigenetiche possibili nei singoli individui, poi per i flussi temporali di variazioni generazionali, poi per tutte le varianti che non hanno avuto luogo ma che avrebbero ben potuto essere (soprattutto queste), e poi chi dice che dobbiamo limitarci a 100 miliardi di partenza, o a considerare solo il nostro Pianeta? Il numero ottenuto non è contenibile in questo universo. Che del resto è ben apeiron già per conto suo.

Il DNA è dunque un'apeiron che abbraccia il mondo e che informazionalmente lo invade. Il DNA è come la Biblioteca di Babele, resa solida e potenzialmente eterna dalla sua capacità di rappresentare tutte le informazioni possibili. Il DNA in linea di principio potrebbe coprire l'universo intero, fagocitare tutto, adattando i propri meccanismi funzionali a tutti gli ambienti nei quali i legami del carbonio sono stabili. E, come sappiamo, la chimica del carbonio, e quella dell'azoto, e dell'idrogeno, e dell'ossigeno, tutte sono simili dappertutto. Da qui in periferia fino al limite estremo dell'universo. Ovunque esso sia.

CAPITOLO II

EPIGENETICA: ESEMPI

1. Il DNA è la nostra vera struttura
2. Imago mundi
3. Esempi estremi del ruolo dell'Epigenetica:
Tempi lunghi: scomparsa e ricomparsa della ali
Tempi brevi: zigoti
Tempi brevissimi: regolazione genica
4. Epigenetica come controllo della complessità

La trasmissione dei caratteri da genitore a discendente attraverso il Genoma è una caratteristica universale della vita, così come la conosciamo. La base fisica dell'evoluzione è proprio questa, ed è tutta qui. Complessa, anche se il suo autogenito inizio deve necessariamente essere stato semplice. Il nostro destino è scritto nel nostro DNA, nella sua sequenza di lettere chimiche, quello che chiamiamo la nostra Genetica; delle quali lettere ognuno di noi può selezionare, ed in parte modificare, l'uso attraverso i meccanismi dell' Epigenetica. "Epigenetica" ha molte definizioni. La più corretta mi sembra essere "Adattamento strutturale di regioni cromosomali tale da registrare, segnalare o perpetuare stati alterati di attività" (Adrian Bird).

1. Il DNA è la nostra vera struttura

Io sono il mio DNA, il mio DNA sono io. Il DNA, nostro primo manuale di comportamento, ci rende liberi e servi allo stesso tempo, individui e cloni, simili e diversi, eterni e mortali, deboli e forti, ha le sue regole severamente

obbligate e allo stesso tempo modulabili per ogni individuo. Primo ed ancestrale insieme di informazioni, il DNA è l'eredità che ho ricevuto e che trasmetterò; a me farne buon uso ed usare al meglio (ma come?) quello che ho ricevuto. Vivendo, usiamo il nostro DNA in due modi.

Il primo consiste nell'usare il DNA proprio come si usa un manuale, che si può leggere per imparare quello che serve, lo stretto necessario, rileggendo le pagine utili molte volte, le stesse pagine tutte le volte che servono, per poi lasciarlo lì sugli scaffali della Biblioteca. L'altro è un modo più creativo, cercando di sfruttarne i piccoli segreti, sforzandoci di leggerlo fino in fondo, cercando di addomesticarlo in qualche modo. Un po' di pazienza: scorriamo quello che segue tenendo ben chiara prima di tutto la distinzione tra uso fisiologico del DNA (per noi stessi, per vivere la nostra vita di individui) ed uso genetico ed epigenetico (per noi, ma soprattutto per le prossime generazioni).

2. Imago mundi

Il DNA è Imago Mundi, ha una sua dimensione storica ben precisa, rispecchia in sé quanto gli è successo nelle generazioni precedenti e lo rivive, lo riflette al di fuori di se stesso, trasmette il suo passato, il suo vissuto. Solo in parte un passato individuale, certamente non un passato astratto. Il mio DNA richiama e assomma il vissuto di tutti gli organismi dai quali provengo, che si sono incontrati ed accoppiati (due a due in atti d'amore) (atti a rimescolare le sequenze), che si sono scambiati informazioni fisiche, che si sono sussurrati parole genetiche e metaboliche che la nostra mente non comprende (ma che, sì, le nostre gonadi capiscono bene), ricombinando le scritture genetiche e la capacità di rispondere all'ambiente, al mondo, alla storia. In termini generali, il DNA è dunque non solo Imago ma anche Istoria Mundi. Se

sapessimo veramente come leggerlo, conosceremmo la nostra storia in modo molto più intimo e sincero, sapremmo veramente chi siamo.

Un buon esempio viene dai dati recenti sulle comparazioni genomiche e paleogenomiche delle migrazioni umane. In breve: abbiamo visto nel capitolo precedente come sia diventato possibile determinare senza molto sforzo né grande spesa la sequenza genomica di un organismo, e farlo più volte, per gruppi. Determiniamo quindi, ad esempio, la sequenza nucleotidica di tutto il genoma di 100 individui che vivono a Trieste, facciamo le dovute medie e correzioni, ed avremo così il genoma triestino medio odierno. Estendiamo a Venezia e alla Croazia, compariamo i risultati, e cominceremo ad avere una prima idea della genomica di quell'angolo di mondo, l'Adriatico del nord-est, oggi. Introduciamo nello studio dati genomici ricavati dall'analisi del DNA da scheletri di due secoli fa trovati in quelle zone, e cominciamo così a delineare nel nostro quadro una dimensione evolutiva di spaccato verticale per dimensione tempo. Ora allarghiamo a tutta l'Europa, poi al mondo, ed estendiamo il tempo di ricerca a mille anni, poi a diecimila, poi a tutto quello che troviamo di conservato ed analizzabile. Avremo allora lo scenario della formazione dei gruppi genetici della specie *Homo*, dei loro spostamenti, delle loro fusioni e delle divergenze. Tutto questo è stato fatto, i dati sono incompleti ed ancora parzialmente insoddisfacenti, ma il quadro d'insieme e le sue regole sono ormai delineate.

A volte i gruppi genetici corrispondono alle culture ed alle lingue, a volte no. Ma questo è in qualche modo atteso e comprensibile. Out-of-Africa ed i vari passaggi per lo Stretto di Bering sono stati confermati, e molto altro. I grandi spostamenti e le grandi fusioni cominciano ad esser chiari. Novità inattese sono frequenti, il fatto ad esempio che ci siano nel sangue dei Sioux più geni simili ai nostri d'Europei che a quelli dei cinesi Han, e così via.

Pian piano il grande romanzo del nostro cammino evolutivo si precisa, la nostra storia esce dall'ombra; del resto la portiamo scritta nel lungo filo doppio avvolto su se stesso nel nucleo cellulare. Ma c'è moltissimo altro, che non immaginiamo nemmeno ma... basta cercare con animo aperto.

3. Esempi estremi del ruolo dell'Epigenetica

Il DNA è la base del sistema vivente e, al tempo stesso, il suo fattore limitante. Io non posso volare, ho difficoltà a correre in alta montagna (bisogna che mi allenino di più), avrò solo 120 anni di vita, il mio DNA è la mia prigione. Il mio DNA, per quanto ottimo, grazie madre, grazie padre, è quanto mi è necessario, ma è anche il mio confine. Inesorabilmente, ogni volta che cerco di spiccare il volo, il mio DNA mi ricorda i miei limiti. Come in una Biblioteca, dove quello che è più importante, quando cerchiamo qualcosa, è quello che in quel momento non c'è. O non c'è più, o non c'è ancora.

Così nel DNA, dove quello che è più importante è cosa abbiamo perduto, oppure cosa c'è ma è silenziato, nascosto nelle pieghe dell'eterocromatina. Quello che potrebbe esserci, o potrebbe esserci stato, ma che ora per me non c'è, o non è disponibile. Oppure è talmente ben nascosto che è come se non ci fosse.

Eccoci allora che stiamo entrando nel vivo dell'Epigenetica, per capire la quale è utile leggere di tre esempi, rappresentativi di tre scale temporali differenti: milioni di anni, mesi, frazioni di secondo. E dato che all'inizio del capitolo abbiamo parlato di volare, per seguire quest'esempio guardiamo un po' nel campo sotto casa, dove abbiamo tutti, o quasi, visto d'estate un insetto secco.