
Il sequenziamento del genoma umano: 20 anni di studi, scoperte e riflessioni

F. Belli

Nel corso del 2022 si sono verificati, in ambito biomedico, due eventi che hanno richiamato l'attenzione tanto della comunità scientifica, quanto dell'opinione pubblica, a livello internazionale, grazie anche alla loro diffusione mass-mediatica. Il primo è stato la pubblicazione, ad Aprile, su *"Science"*, dei risultati ottenuti da un Consorzio Internazionale che ha sequenziato per la prima volta al 100% un genoma umano, colmando tutte le lacune delle indagini precedenti; il secondo, ad Ottobre, riguarda l'assegnazione del premio Nobel per la Medicina e la Fisiologia a Svante Pääbo, pioniere e tra i padri dell'antropologia molecolare e della paleogenetica, sì che discipline e premiato hanno visto riconosciuta l'importanza di ricerche, dalle ricadute dottrinali e pratiche anche nella medicina moderna.

Obiettivo iniziale di questa conferenza era quello di illustrare i progressi compiuti negli ultimi anni riguardo il sequenziamento dei genomi umani attuali e antichi e delle altre specie ominine con cui abbiamo convissuto e ci siamo ibridati: un aggiornamento del nostro precedente convegno in Accademia del 2013. I due importanti eventi ricordati sopra hanno corroborato ulteriormente il progetto; conferenza, e memoria che presentiamo, si articoleranno in due parti: illustrazione del primo genoma umano completamente sequenziato, tecnologia e risultati innovativi già ottenuti; aggiornamento riguardo i confronti fra genomi antichi e moderni,

sottolineandone le ripercussioni in medicina e clinica pratica, nonché gli apporti di geni esogeni o "alieni" nel DNA umano, avvenuti sia dopo incroci con altre specie ominine, che a seguito di passaggi orizzontali unidirezionali e integrazione di geni microbici, materia quest'ultima ancora largamente inesplorata.

Il sequenziamento del genoma umano: dai primi risultati parziali di inizio secolo, al completamento oggi raggiunto

Nel febbraio 2001 apparve su *"Nature"* il frutto del lungo, e costosissimo, lavoro dello *"Human Genome Project"* che, onestamente, lo pubblicava come sequenziamento iniziale e parziale del genoma umano¹: ma il battage pubblicitario fu tale che venne presentato, e tanti lo percepirono come tale, quale risultato completo e definitivo. Ci sono voluti vent'anni per giungere ad un primo sequenziamento completo, anni nel corso dei quali diversi gruppi interdisciplinari (l'impresa al momento non è attuabile da singoli o pochi ricercatori per la sua complessità) si sono dedicati a colmare le numerose lacune presenti nella conoscenza del nostro genoma, strutturali e, di conseguenza, funzionali, ognuno aggiungendo comunque tasselli importanti. Ricordiamo fra gli altri il *"Genome Reference Consortium"* (GRC, 2013), il cui genoma sequenziato nove anni fa, GRCh38, è stato assunto come riferimento fino ad oggi. Il gruppo che ha annunciato sulle pagine di *"Science"* il primo sequenziamento completo di un genoma umano è il *"Telomere-to-Telomere Consortium"*², che propone come nuovo

riferimento il proprio lavoro, siglato T2T-CHM13v1.1.

Come nei precedenti genomi di riferimento, anche il lavoro di “T2T Consortium” si è basato sul sequenziamento di tutti i cromosomi, compreso X, di una linea cellulare prelevata da una mole idatiforme (CHM), “Complete Hydatidiform Mole”. La maggior parte dei genomi CHM sono omozigoti con un cariotipo 46,XX. Analisi ancestrali evidenziano che la maggior parte del genoma CHM13 è di origine europea, comprese regioni derivate da introgressioni neanderthaliane.

Alla realizzazione del progetto del “T2T-Consortium” hanno lavorato un centinaio di ricercatori di tutto il mondo, perlopiù biologi, genetisti, bioinformatici; fanno capo a 54 istituzioni e centri di ricerca tra i più prestigiosi nei campi delle scienze biomediche, genetica di base e applicata, biotecnologie, informatica e procedure computazionali.

Se assumiamo un semplice parametro, il numero di coppie di basi (bp) analizzate e sequenziate, quale indicatore delle lacune colmate nel corso degli anni nei genomi di riferimento proposti dai diversi gruppi di lavoro, osserviamo come si è passati dai 2.580.000.000 (82.7%) bp del 2001, “Human Genome Project”, ai 2.800.000.000 (89.8%) bp fra gli anni 2003/2010, ad una quota attorno ai 2.900.000.000 bp (92%) del GRCh38 del 2013. Il “T2T-Consortium” ha colmato l’8% rimanente, per cui ha presentato il suo risultato con lo slogan: 3.117.275.501 bp sequenziate, 0 omissioni o lacune. Ma, al di là dei numeri, l’importanza del lavoro svolto per ottenere T2T-CHM13v1.1, è nella natura e nel significato strutturale e funzionale delle sequenze che costituiscono l’8% rimasto finora nel “dark side of our genome”, distribuite in quote variabili praticamente in tutti i cromosomi. Ecco dunque quali sono i progressi e le novità raggiunte:

- Identificazione e sequenziamento completo delle oltre 3 miliardi bp (in precedenza ne mancavano almeno 200 milioni).
- Decifrazione delle parti più difficili da sequenziare: centromeri e telomeri dei cromosomi, bracci corti dei 5 cromosomi acrocentrici e relativi satelliti, elementi da cui dipendono, durante e dopo la replicazione del DNA, ordinamento e corretta separazione di tutto il materiale genetico di un cromosoma, nonché sequenze e geni regolatori di percorsi metabolici che interferiscono con la durata di vita della cellula. Sono tutti caratterizzati da sequenze altamente ripetute, che hanno finora reso complicato il loro corretto sequenziamento.
- Sequenziamento di numerosissime sequenze riferibili a DNA ribosomiale (rDNA).
- Individuazione di 2000 nuovi geni, molti disattivati, ma altri attivi o potenzialmente attivabili; 2 milioni di nuove varianti geniche sconosciute.
- Sequenziamento completo, senza lacune ed errori, di sequenze dapprima analizzate in modo approssimativo: sono emerse 622 nuove varianti in geni di cui è noto l’impatto e il collegamento con patologie e quadri clinici.
- Sequenziamento completo del DNA mitocondriale: 16569 bp e 37 geni.

Manca, nel lavoro di “T2T-Consortium”, per scelta tecnologica e procedurale, il sequenziamento completo del cromosoma Y³, obiettivo peraltro già raggiunto da altri gruppi.

Il numero di geni codificanti si è attestato a quota 19.969, 140 dei quali totalmente nuovi; fra questi, 4 sono indicati di interesse clinico-medico (CFHR1, CFHR3, OR51A2, UGT2B28): hanno caratteristiche in comune, il differente grado di presenza e attivazione in rapporto al numero di copie, la proprietà di codificare per recettori delle cellule T verso

immunoglobuline. I 140 geni totalmente nuovi ed esclusivi di T2T-CHM13 sono perlopiù localizzati nelle regioni pericentromeriche e nei bracci corti dei cromosomi acrocentrici. Vi sono tuttavia oltre 63.000 sequenze, il 6% in più rispetto ai genomi precedenti, che devono essere funzionalmente analizzate: alcune potrebbero identificarsi in nuovi geni codificanti. Come segnalato, sono state individuate numerose sequenze ripetute (33% in più), e sono stati sequenziati tutti i satelliti dei cromosomi acrocentrici. Il nuovo genoma aggiunge o corregge 238 Mbp di sequenze che non erano allineate correttamente in GRCh38 (> 1%), sequenze che comprendono i satelliti centromerici (76%), duplicazioni segmentali (19%), rDNA (4%); di queste, 182 Mbp non hanno un allineamento primario (non si ritrovano) in GRCh38 e sono esclusivi di T2T-CHM13. Come risultato finale, T2T-CHM13 aumenta il numero di geni conosciuti nel genoma umano.

Confrontando “vecchi” e nuovo genoma, T2T-CHM13 è non solo più completo, ma anche più preciso e rappresentativo e referente di tutte le varianti brevi e lunghe emerse anche nei campioni di genomi umani ancestrali; riduce contemporaneamente sia i falsi positivi che i falsi negativi tra le varianti, generatisi in GRCh38 a causa dell’aggiunta a posteriori di sequenze mancanti o improprie, o per l’esclusione di altre o false duplicazioni. Questi miglioramenti, anche per la riduzione prossima allo 0 di sequenze rivelatesi incongrue o errate, hanno permesso di eliminare migliaia di varianti spurie. Infine, il “nuovo” genoma si è dimostrato più rappresentativo delle “*human copy number variations*” rispetto a GRCh38 e, se confrontato con 268 genomi umani del progetto SGDP (“*Simons Genome Diversity Project*”), è 9 volte più predittivo di questi, sempre riguardo a hCNV. Questi risultati sottolineano sia la qualità dell’assemblaggio che la stabilità genomica della linea cellulare da cui è derivato.

Rimandiamo al report edito da “*Science*” per visualizzare le vecchie e nuove sequenze e i cromosomi interamente sequenziati e assemblati, mediante un nuovo programma che automaticamente, dopo il sequenziamento, riporta in un grafico bidimensionale l’assemblaggio a “*string graph*” ad alta risoluzione e l’allineamento di ogni singolo cromosoma, bracci corti e lunghi, con particolare riferimento ai bracci corti dei 5 cromosomi acrocentrici e relativi satelliti, alle regioni centromeriche e telomeriche di tutti i cromosomi, al rDNA, cioè tutte le sequenze finora rimaste escluse da un completo e corretto sequenziamento; soprattutto si conferma la frequenza di sequenze altamente ripetute.

Oltre i 5 cromosomi acrocentrici (13,14,15,21,22), presentavano ampie lacune, ora colmate, i cromosomi 1,9,16: anche questi sono caratterizzati da diverse sequenze ripetute. T2T-CHM13 manifesta un’origine prevalentemente europea, e conferma la presenza di regioni derivate da introgressioni neanderthaliane.

L’assemblaggio e l’allineamento a “*string graph*” ad alta risoluzione del genoma CHM13 permette di dipanare regioni con sequenze ripetute e/o ad alta similitudine, come i bracci corti dei cromosomi acrocentrici: finora avevamo, quale rappresentazione grafica, una matassa confusa e intricata, ora le regioni sequenziate al 100% sono correttamente assegnate a ciascun cromosoma, comprese quelle appartenenti a rDNA. Tutto questo, come vedremo, è la base fondamentale per assegnare a ciascuna sequenza genica posizionamento esatto e ruolo, operazione che prende il nome di “annotazione del genoma”.

In definitiva, per ogni cromosoma, bracci lunghi e corti sono mostrati con i geni “annotati”, la percentuale di unità geniche metilate, e l’annotazione dei satelliti e delle ripetizioni è basata su un codice, ad es. di colori, per differenziarli. Le sequenze di rDNA

sono illustrate con il numero di copie, caratterizzate da una elevata similarità fra di loro, situazione che ne ha impedito il mappaggio con le procedure convenzionali. Le “mappe di calore” (*heatmaps*) evidenziano un’alta identità dei bracci corti di ciascun cromosoma acrocentrico con quelli degli altri 4. I bracci corti distali comprendono strutture satellite conservate e ripetizioni invertite, mentre i bracci corti prossimali mostrano una diversità strutturale.

Un esempio di nuovi dati, applicabili in clinica, individuati nel genoma T2T-CHM13, è rappresentato dalla genetica della distrofia muscolare fascio-scapolo-omerale (FSHD). È stata studiata una regione duplicata del cromosoma 4, quella subteleromica 4q, che è associata alla malattia; questa regione, molto complessa, comprende geni implicati nell’eziologia della malattia: FRG1, FSHD region gene 1, FRG2, FSHD region gene 2, e una ripetizione macrosatellite D4Z4 contenente un doppio gene homeobox 4 (regolatore e conservato) DUX4. Numerose duplicazioni di questa regione hanno complicato in precedenza l’analisi genetica della FSHD.

L’assemblaggio di T2T-CHM13 rivela 23 *paralogs* (geni o tratti di DNA appartenenti a un gruppo di sequenze omologhe all’interno della stessa specie) di FRG1 attraverso tutti i cromosomi acrocentrici e i cromosomi 9 e 20. Questo gene ha subito una recente amplificazione nei primati. Comunque, solo 9 FRG1 *paralogs* erano stati trovati in GRCh38, a causa delle difficoltà di un’analisi basata sulle sequenze. Fra questi 9, uno, FRG1DP, è localizzato nella regione centromerica del cromosoma 20 e presenta un’alta identità, 97%, con diversi altri paraloghi, FRG1BP4/10. Con il mappaggio attuale, è stato inoltre dimostrato come nel vecchio genoma fossero emerse molte varianti falsamente positive, ora eliminate. Molti paraloghi FRG1 sono presenti anche in altri genomi umani sequenziati, e tutti, tranne FRG1KP2/3, hanno a monte isole CpG e sono fortemente

espressi in CHM13. Pertanto, oggi sappiamo che alla FSHD sono associati paraloghi e varianti ben più numerosi di quanto evidenziato dai precedenti, incompleti genomi.

Per quanto riguarda il cromosoma Y, le scoperte più recenti, che sintetizziamo per completezza, sembrano quasi ridimensionarne il ruolo⁴. La trasformazione delle ovaie in testicoli e l’input alla produzione di testosterone non sono dovute solo all’azione dei geni SRY, ma anche dei geni Sox9 e Enh13, che si trovano sul cromosoma 17. Quasi la metà (ma il numero è in aumento) degli uomini >70 anni perde il cromosoma Y dalle cellule staminali emopoietiche e di altri fenotipi: le cellule figlie non lo ricevono più; è stato dimostrato che questi soggetti sviluppano più frequentemente tumori, patologie cardiache e vivono circa 6 anni di meno degli uomini che conservano Y. I maschi Y- vanno incontro a processi fibrotici a carico di vari organi, cuore compreso, per l’attivazione del sistema fibrogenetico indotto da TGFβ: la relazione con la perdita di Y non è chiara. In ogni caso, con l’età l’espressione dei geni del cromosoma Y si riduce, le proteine codificate diminuiscono e le funzioni correlate si affievoliscono. I biologi evolucionisti evocano persino la progressiva scomparsa di Y nella nostra specie, evento già osservato in altre specie animali.

Per concludere, riportiamo alcuni numeri relativi al nostro genoma e un confronto con altre specie viventi⁵. Prima del genoma sequenziato nell’ambito dello «*Human Genome Project*», < 2000, il numero di geni presenti in cDNA era stimato fra 40.000 e 100.000; con i primi sequenziamenti (2001-2012) il numero si è ridimensionato a 26.000 – 30.000. I genomi di riferimento analizzati fra il 2013 e il 2020 (GRCh38, ad es.) hanno riportato quote comprese tra 19.890 e 20.700. T2T-CHM13 ne ha depennati, in base ai dati della proteomica e della trascrittomica, ulteriori 1700 la cui azione

non si traduce in codifica di proteine, ma ne ha aggiunti un centinaio «nuovi», rintracciati solo in quest'ultimo riferimento, per cui la cifra ultima di geni codificanti nucleari è 19.969; a questi vanno aggiunti 37 geni presenti nel DNA mitocondriale. Nel genoma vi sono però diverse migliaia di geni nc, pseudogeni, misconosciuti, totalmente inattivi o comunque ancora di difficile decifrazione. Il 90% dei geni umani ha un'origine antichissima, è condiviso o è stato ereditato dagli animali a simmetria bilaterale; il rimanente 10% è più recente, da 50 milioni *ya* (*years ago*, anni fa) a oggi, dopo la comparsa dei primati, con i quali è condiviso.

Classi e ordini di animali presentano dimensioni del genoma assai diversificate; così nei mammiferi il numero di bp varia da 1.5 a poco più di 4 miliardi: come abbiamo visto, quello umano è di 3.1 miliardi bp, molto simili i genomi di cane e gatto, 3.6 miliardi bp il DNA della grande megattera norvegese⁶. Le dimensioni dei genomi dei rettili sono molto simili a quelle dei mammiferi, negli uccelli non superano i 2.5 miliardi bp, molto variabili nelle altre classi di animali i valori riscontrati: i pesci cartilaginei fra 2.3 e 8 miliardi bp; i molluschi da 0.5 a 2.8 miliardi bp; gli insetti da poche migliaia a > 10 miliardi; infine, anfibi e pesci ossei da 0.5 a > 100 miliardi bp! E, tra questi ultimi, ricordiamo *Protopterus Aethiopicus*, un antico pesce osseo dei mari africani il cui genoma > 133 miliardi bp! Legge comune è che quanto più un genoma è grande, tanto più contiene sequenze ripetute anche milioni di volte. Le dimensioni di un genoma, invece, non correlano né con il numero di geni presenti, né con la complessità e le proprietà funzionali della specie.

Brevi note tecniche

In realtà il sequenziamento è solo il primo passo di una serie di procedure che permettono di conoscere, di una struttura complessa, come un genoma a DNA e/o RNA, natura biochimica e molecolare dei singoli componenti, allineamento, configurazione e

rapporti: seguono altri steps ugualmente importanti che completano un iter chimico e informatico, ormai quasi completamente automatizzato, multiprocedurale e multidisciplinare.

SEQUENZIAMENTO: processo attraverso cui si determina l'esatta sequenza dei nucleotidi che costituiscono l'intero genoma di un organismo vivente; oggi è notevolmente facilitato dalla rapida evoluzione delle tecnologie automatizzate e di quelle bioinformatiche in relazione all'acquisizione, memorizzazione, analisi e interpretazione dei dati. Si articola in varie fasi: ESTRAZIONE, PURIFICAZIONE, ISOLAMENTO del DNA, la sua frammentazione casuale e il clonaggio dei frammenti ottenuti in vettori stabili alla propagazione in cellule di batteri o di lievito; le attuali tecnologie consentono il sequenziamento fino a 1.000.000 di nucleotidi per ogni singola azione. I vettori utilizzati sono oggi i cromosomi batterici artificiali (BAC, *Bacterial Artificial Chromosome*) che possono contenere fino a 300 kb. I cloni così ottenuti vengono ordinati a seconda della loro localizzazione nel contesto della porzione cromosomica di provenienza, producendo la mappa fisica di questa parte di genoma. Dopo un'ulteriore frammentazione di ciascun frammento clonato, si procede al sequenziamento nucleotidico vero e proprio: si utilizza perlopiù la tecnica enzimatica, più adatta al sequenziamento automatico, sviluppata da F. Sanger.

Per ricostruire la sequenza completa di un genoma si fa quindi ricorso a una strategia, chiamata *shotgun*, che consiste nell'ASSEMBLAGGIO e nell'ALLINEAMENTO delle migliaia di piccoli frammenti del genoma stesso, provenienti da librerie plasmidiche, dei quali si conosce la sequenza, usando un sistema informatico e software dedicati per cercare i loro punti di attacco e sovrapposizione nucleotidica (VALIDAZIONE). I frammenti contigui (*contigs*) assemblati in

modo completo coincidono con i cromosomi del genoma.

Seguono la PULIZIA: eliminazione di sequenze incongrue, errate, non coincidenti e sovrapposte e l'ANNOTAZIONE DEL DNA: il processo per identificare le posizioni dei geni e tutte le regioni codificanti in un genoma e determinare cosa fanno quei geni.

I metodi di sequenziamento, ma soprattutto clonazione, assemblaggio e validazione convenzionali hanno presentato dei limiti, sfociati nel mancato obiettivo del 100% di sequenziamento completo del genoma umano. "T2T-Consortium" ha superato questi limiti implementando le tecnologie precedenti e creandone di nuove, per tutta la procedura: "PacBio's circular consensus sequencing" (HiFi), "Oxford Nanopore ultralong-read sequencing" (ONT), "Illumina PCR-Free sequencing" (ILMN). Il principio dell'assemblaggio di T2T-CHM13 si basa sulla rappresentazione grafica "assembly string graph" (grafico a corde), costruita direttamente dalla tecnologia "HiFi", che opera anche la validazione e il polishing, in un "string graph" bidirezionale.

Alcuni risultati e punti fermi raggiunti

Delegittimazione del concetto di razza. Antropologia molecolare e genetica hanno dimostrato in maniera definitiva che la classificazione e la suddivisione della specie umana in razze o sottogruppi è inconsistente, la scienza ha decostruito ogni possibile tentativo di farlo. La nostra specie è universalmente unica pur in una complessità variegata. Come già aveva notato Darwin, *"...le differenze fra le popolazioni, anche se vistose, sono perlopiù irrilevanti, mentre vi è una grande uniformità nelle caratteristiche veramente importanti, comprese quelle mentali e cognitive: nonostante le differenze esteriori tra aborigeni americani, negri africani ed europei, quello che veramente colpisce, dai tanti aspetti del carattere, è quanto le loro menti siano simili tra loro..."*.

L'antropologo ed etnologo C. Levi-Strauss⁷ e il genetista L.L. Cavalli-Sforza⁸ hanno scritto libri fondamentali che bocciano scientificamente l'esistenza delle razze. R. Levi-Montalcini⁹ ha affermato: *"...la continua variazione della specie umana è stata plasmata dagli incessanti spostamenti e dalle relazioni con gli ecosistemi ed è sempre stata alla base della nostra stessa evoluzione..."*.

Malattie mono/poligeniche, epigenetiche e multifattoriali.

La conoscenza più approfondita del nostro genoma conferma e ribadisce che:

1. Le malattie da alterazione, assenza o ipo-espressione di un solo gene sono complessivamente numerose, diverse migliaia, ma ognuna rimane confinata perlopiù nel campo delle rarità. Oggi conosciamo meglio non solo quelle da deficit di geni nucleari, ma anche mitocondriali. In Italia le malattie "rare" monogeniche incidono con una frequenza complessiva di 42 casi/100.000 abitanti, i pazienti sono attualmente 255.000 e fanno capo a 300 presidi specialistici¹⁰.
2. Le malattie la cui eziologia dipende dal concorso di più geni sono ben più numerose, ma forse ancor meno inquadrabili e classificabili, dunque eziologicamente nebulose: i geni coinvolti sono sì ben noti, ma peso e ruolo nel determinismo dei diversi quadri clinico-patologici variano da forma a forma, da soggetto a soggetto, configurandosi pertanto una miriade di situazioni diversificate.
3. Cominciamo a conoscere patologie la cui eziologia è legata anche a fattori epigenetici, ai meccanismi che condizionano il grado di espressività e attivazione/repressione di molti geni, lungo le sequenze di DNA.
4. Ma la stragrande maggioranza delle patologie, neoplastiche, degenerative, dismetaboliche, nasce dal concorso di fattori endogeno-genetici ed esogeni,

infettivi, inquinanti, iatrogeni, talora anche con l'intervento di meccanismi epigenetici: una multifattorialità che rende ancora ardua l'esatta comprensione della relazione cause/effetti.

Progressi in neuroscienze. È questo il settore cui la genetica ha negli ultimi anni offerto il maggior contributo di conoscenze, sia nella morfo-elettrofisiologia che nelle ricadute cliniche. I rapporti fra GENI E SVILUPPO CEREBRALE, crescita, maturazione e modellamento del cervello umano hanno basi genetiche ben precise, che si esprimono durante i primi, decisivi anni di vita; l'uomo è il solo essere vivente in cui il proprio cervello impiega (quasi) vent'anni per maturarsi compiutamente. La lunghezza dell'infanzia e dell'adolescenza, con il perdurare prolungato di tratti giovanili, fa sì che acquisiamo e immagazziniamo esperienze, che diventano decisive e fondamentali nella maturazione cognitiva. I circuiti neurali nel corso dello sviluppo subiscono cambiamenti strutturali e funzionali, come fossero «scolpiti» dalle esperienze che facciamo. Durante l'accrescimento, i cambiamenti morfo-funzionali del cervello sono condizionati dall'attività di geni la cui trascrizione e codifica è regolata a loro volta dalla stessa attività dei neuroni: si crea così un meccanismo di feed-back in cui diversi protagonisti molecolari e cellulari si autoregolano vicendevolmente. Alla base della plasticità e delle caratteristiche uniche del nostro cervello vi sono dunque geni, in un contesto di lungo accrescimento, acquisizione di esperienze e maturazione, tutti fattori che interagiscono fra loro. Riportiamo alcuni dei geni coinvolti nello sviluppo cerebrale, con particolare attenzione a quelli studiati di recente:

- Gene regolatore HARE5: controlla l'attività di altri geni coinvolti nello sviluppo della corteccia cerebrale; la variante umana permette una proliferazione delle cellule corticali di gran lunga superiore a quella degli altri primati.

- ARHGAP11B: fondamentale nello sviluppo embrionale della neocorteccia, è specifico dei rappresentanti del genere *Homo*, mentre è assente negli altri primati; insieme ad una cinquantina di altri geni, regola la proliferazione delle cellule progenitrici dei neuroni e della glia e del ripiegamento degli strati corticali, da cui dipende l'estensione della corteccia stessa.
- Altre cellule, di recente individuazione, che nell'embrione umano proliferano ad un tasso nettamente superiore a quello di ogni altro animale, sono le «staminali neurali della glia radiale esterna oRG»¹¹: sotto controllo genico, permettono l'eccezionale sviluppo di aree corticali critiche per funzioni cognitive specifiche dell'uomo.
- GENE per OSTEOCRINA e il suo PROMOTER, MRE: ad azione ritardata ma protratta, si trova nella neocorteccia umana, soprattutto nei neuroni negli anni della crescita e dello sviluppo infantile, regola la forma dei dendriti, promuove l'allungamento degli assoni e altri cambiamenti morfo-funzionali dei neuroni neocorticali, che acquisiscono le caratteristiche proprie della specie umana¹². Negli altri vertebrati regola la crescita di ossa e muscoli: è stato definito «un gene riciclato», esempio di exaptation o cooptazione funzionale.
- Geni specifici dell'uomo, assenti negli altri primati: ZP2, iperespresso nello striato e nel cervelletto nei primi anni di vita¹³. TH: presiede alla produzione di dopamina, fortemente espresso nella neocorteccia e nello striato. MET: fortemente espresso nella corteccia prefrontale, è un gene connesso, fisiologicamente, con funzioni cognitive superiori e peculiari umane, ma patologicamente con lo spettro autistico. Ma i più importanti sono i GENI NOTCH2NL^{14, 15}: fanno parte di un sistema di famiglie geniche, correlate tra loro, presenti in forma attiva nella

corteccia umana in accrescimento; ritardano il differenziamento delle cellule staminali in cellule nervose corticali, accumulando le prime che evolvono in un maggior numero di neuroni maturi. Si trovano in una regione del genoma, 1q21.1, nota per ospitare altri geni coinvolti in disordini neurologici e in alcune forme di autismo, nonché in difetti congeniti caratterizzati da macro o micro-encefalia.

- NOVA1: controlla l'espressione di molti altri geni durante lo sviluppo precoce del cervello.
- SRGAP2C: aumenta la densità delle connessioni neurali.

Molti neuroscienziati si dedicano a cercare eventuali basi genetiche, o quantomeno un coinvolgimento di geni, in aspetti precipi della nostra specie, quali autocoscienza, creatività e socializzazione, geni che sarebbero pertanto assenti in tutti gli altri primati e ominidi estinti; ben 300 geni sono stati in diversi studi indicati o "sospettati", ma siamo ben lontani da una dimostrazione plausibile¹⁶. Come ha scritto l'antropologo Ian Tattersall: "*...Il nostro comportamento non è fisso o determinato dai geni; ma è emerso per la necessità di adattarsi rapidamente a condizioni difficili e per comunicare...*"¹⁷.

Alcune mutazioni di geni coinvolti nello sviluppo evolutivo del cervello, si sono rivelate un'arma a doppio taglio: vantaggiose in età giovanile, stimolando una maggior crescita dei neuroni e delle sinapsi, dannose per gli anziani in quanto connesse con diverse patologie; una sorta di contropartita evolutiva di determinati adattamenti. In particolare, fra gli *enhancer*, sequenze di DNA "driver" di geni vicini, ne sono stati individuati 90, attivi e specifici nei neuroni umani, profondamente diversi rispetto a quelli delle scimmie antropomorfe: queste sequenze regolano geni prossimali implicati nella crescita cerebrale, ma sono anche coinvolte, nell'anziano, nella patogenesi di

neuropatie degenerative come l'Alzheimer. Al contrario sono state individuate altre sequenze che, pare, proteggano dall'Alzheimer, ma frenano l'attività di alcuni geni della crescita in età giovanile. È il paradosso del doppio ruolo evolutivo dei neurogeni: favoriscono lo sviluppo cerebrale fra i giovani, ma devono pagare un tributo all'evoluzione, aumentando il rischio di malattie degenerative col crescere dell'età. Le principali e più frequenti malattie neurodegenerative, come l'Alzheimer, dello sviluppo neurologico coinvolgenti linguaggio e comunicazione sociale, come lo spettro autistico, psichiatriche, quali la schizofrenia, sono dovute al concorso di mutazioni poligeniche, eventi epigenetici e fattori esogeni perlopiù ancor oggi non conosciuti. Riportiamo a titolo esemplificativo alcuni geni individuati in 4 patologie neuropsichiche:

- MALATTIA DI Alzheimer¹⁸⁻²². Geni XRCC1 e PARP: implicati nei meccanismi di riparazione del DNA. Varianti APOE2/3/4: maggior o minor capacità di degradare β -amiloide. Mutazione TREM2/R47H: < funzionalità della microglia. Gene APP: inibisce la generazione dei peptidi β -amiloide. CD33, CR1, BIN1, MS4A6A: relativamente frequenti, ma con ruolo ancora nebuloso. Mutazioni geni metabolismo proteina TAU.
- SPETTRO AUTISTICO²³. CNV o SNP'S in circa 800 geni coinvolti nell'adesione tra neuroni (NLGN1, ASTN2), nella formazione di sinapsi (SHANK2, SYNGAP1, AP2, DX53-PTCHD1) e nella sintesi di ubiquitina (UBE3A, PARK2, RFW2, FBO40). 45 geni comuni ad autismo, ADHD, dislessia e schizofrenia, che codificano per proteine indispensabili nella crescita dei nevriti e nella comunicazione inter-neuronale. Mutazioni SETD5 e BCKDK, comuni ad autismo e forme varie di ritardo mentale.
- SLA²⁴. La malattia può essere divisa, in base alle anomalie genetiche, in 4 fasi

progressive; le mutazioni geniche all'esordio e nelle fasi tardive sono diverse, così come diversi i gradi di espressione genica nelle cellule midollari. Una mutazione frequente è a carico del gene C9ORF72, nei malati e familiari di SLA, associata a meccanismi di danno dei motoneuroni.

- SCHIZOFRENIA^{25, 26}. Numerosi sono i geni presumibilmente correlabili con la sua eziopatogenesi, nessuno singolarmente determinante: siamo pertanto ancora lontani da un trattamento che abbia come target cause e/o concause, su base genica, della malattia; la quale appare come una situazione in cui più geni hanno globalmente un ruolo intricato, ma singolarmente sfumato, nell'eziopatogenesi e nello sviluppo dei sintomi, oltre fattori esogeni. Sono coinvolte mutazioni C4 (sfolgimento sinaptico), GRM3 (signalling), DRD2 (dopamina).

Invecchiamento e longevità: esistono geni o marcatori biologici individuali o di popolazione²⁷⁻²⁹?

L'uomo è il primate più longevo. Abbiamo geni che permettono di controllare infezioni, infiammazioni e parametri metabolici, quali il colesterolo, acquisiti durante la storia evolutiva: APOE. Gli stessi geni, nel tempo, sono però divenuti in parte anche fattori di suscettibilità per malattie croniche, autoimmuni, dismetaboliche, neoplastiche. Geni etichettati "di longevità" o "anti-età" ne sono stati individuati diversi, dimostrati, in maniera definitiva, nessuno: ecco alcuni esempi:

- TERC e PAI-1: regolano la lunghezza dei telomeri.
- Varianti geniche studiate in popolazioni longeve (Ecuador, Sardegna, Giappone), come gli antagonisti o i regolatori di alcuni *Growth Factor*; PAI-1 (Amish), anti-trombotico.
- Geni che regolano nelle cellule staminali emopoietiche quantità, tipologia e

velocità di differenziazione e replicazione.

- mTOR e Pol-III: interferiscono con la crescita e l'invecchiamento cellulare e il risparmio calorico ed energetico.
- STK17A, COA1 e altri geni implicati nella riparazione del DNA.

La nostra longevità non è solo un evento biologico, ma dipende anche da situazioni socio-economiche, parentali, organizzative, nutrizionali e dallo stile di vita; cercare geni che in assoluto portano (e in buona salute) ad un'età avanzata è fuorviante: l'elisir di lunga vita non è (o non è solo) nel DNA. La longevità non è scritta nel DNA³⁰.

I grandi progetti di studio internazionali sul genoma umano (e non solo...)

Il sequenziamento completo del genoma umano è sì un punto di arrivo di anni di studi e applicazioni tecnologiche, ma costituisce lo start per nuovi progetti scientifici, che si presentano affascinanti e allo stesso tempo ancor più complessi da un punto di vista organizzativo e metodologico. Ecco che nascono consorzi, enti internazionali e multidisciplinari, per gestire programmi di ricerca sempre più ambiziosi. Nell'immediato e nel prossimo futuro, l'attenzione dei ricercatori si sposterà gradualmente dall'indagine meramente strutturale di geni e sequenze, al dato funzionale e ai complessi meccanismi di regolazione di singoli o gruppi di geni, al grado di espressione in condizioni fisio-patologiche, alle differenze fra cito e istotipi, alle relazioni epigenetiche. Tra i progetti già operativi ricordiamo³¹⁻³⁵: «*HUMAN KNOCKOUT PROJECT*», mappatura delle funzioni di geni e delle conseguenze fisio/patologiche della loro assenza o inattivazione; «*ROADMAP EPIGENOMICS PROJECT*», mappatura epigenetica del genoma; «*GENOTYPE TISSUE EXPRESSION*», mappatura dell'espressione dei geni umani, la loro regolazione e le differenze nei diversi tessuti; «*ALPHA FOLD PROTEIN STRUCTURE DATABASE*», algoritmo che ha sviluppato la struttura 3D di 200 milioni di proteine.

I geni raccontano la nostra storia. Antropologia molecolare e genetica svelano il percorso di H.sapiens. Aggiornamento della conferenza del 2013

A seguito del sequenziamento dei genomi umani antichi e attuali, di altre specie del genere *Homo* estinte, di primati con cui confrontarci, negli ultimi anni, grazie anche alle nuovissime metodologie d'indagine, abbiamo raccolto una miniera di nuovi, e spesso rivoluzionari, dati, grazie a tanti gruppi di ricerca, come quello di Svante Pääbo, ma non solo; in questa seconda parte della nostra memoria, aggiorniamo quanto presentammo in Accademia nel 2013: recenti studi e scoperte che stimolano diverse riflessioni.

La conferenza di cui sopra verteva su una serie di argomenti e applicazioni, che qui ricordiamo^{36, 37}. Ricostruzione della storia evolutiva della nostra specie e confronto con i genomi di altri *Homo* e di primati; percorsi geografici di *H.sapiens*: migrazioni e colonizzazione della terra; geni e fisiologia: come si è evoluto il linguaggio; coevoluzione geni-cultura: l'esempio della lattasi; malattie umane antiche e nuove: evoluzione e cambiamenti nelle interazioni fra patogeni ed ospite; ricostruzione di antiche vicende storiche; ricostruzione di antiche usanze e percorsi culturali. E ricordiamo anche quanto disse S. Pääbo, una sorta di manifesto programmatico dell'antropologia molecolare: *"Il genoma è non solo contenitore di tutta l'informazione genetica, ma museo esso stesso del passato dell'uomo, delle specie ancestrali, un museo che, imparando a percorrere e a decodificare, può rivelare tantissimo della nostra storia biologica e mentale"*.

Quanto è antica la nostra specie? Il numero di varianti geniche nelle sequenze di mt/nDNA che separa due popolazioni è direttamente proporzionale al tempo a partire dal quale esse hanno condiviso lo stesso antenato: quanto più a lungo sono state separate, tanto maggiore sarà

l'accumulo di mutazioni casuali con cui i geni si evolvono.

Questo accumulo è definito tasso di evoluzione; per mtDNA è di 1-2 mutazioni puntiformi/100 nucleotidi/milione di anni. mtDNA è stato assunto come un "orologio molecolare" che batte con un ticchettio molto veloce e costante nel tempo, in termini di evoluzione, e che si presta a documentare gli studi di archeo-genetica e antropologia molecolare: l'evoluzione di specie che si sono diversificate o di popolazioni umane che si sono differenziate. Il tasso di mutazione per nDNA risente di numerose variabili e si presta ad errori e controversie metodologiche, oggi in gran parte risolte per cui anch'esso è entrato nel computo del tempo dell'orologio molecolare. Questo termine fu adottato nel 1962 da E. Zuckerkandl e L. Pauling, che si avvalsero delle mutazioni riguardanti l'emoglobina; successivamente vennero utilizzate altre proteine come le immunoglobuline, infine i marcatori genetici.

Dati archeologici, antropologici e genetici confermano nel porre la divaricazione ultima fra Hominini-pre umani e Hominini non-umani 6.500.000/7.000.000 *ya*: nessun fossile anteriore, a tutt'oggi, può essere inserito tra gli antenati o precursori di generi e specie da cui ha tratto origine la nostra. Ma non sempre i tre filoni di studio hanno fornito dati concordanti: lacune e divergenze vistose hanno marcato la ricostruzione temporale del nostro passato, e solo di recente si è arrivati ad una sintesi più attendibile e precisa grazie a nuove scoperte: eccone in breve la storia.

Fino a pochi anni fa, i fossili più antichi recuperati della nostra specie erano quelli di Omo Kibish, 195.000 *ya* e Herto, 164.000 *ya*; a livello genetico, l'orologio molecolare del nostro genoma indica in 200.000 *ya* l'inizio della sua specificità, mentre il confronto fra mtDNA umano e neanderthaliano dimostra una sostanziale non identità. Più di recente, nDNA antichi e moderni, umani e non, e nuovi dati dell'orologio molecolare hanno

indicato per *H.sapiens* un'origine più lontana. Anche le scoperte archeologiche hanno sollevato molti interrogativi: manufatti complessi, polimorfi, poliuso, frutto di menti cognitivamente evolute, sono stati datati > 250.000 *ya*, quindi o artefici furono specie diverse da *H.sapiens*, o quest'ultimo è comparso prima di quanto i fossili disponibili avessero finora indicato. La scossa clamorosa è venuta con la datazione dei fossili scoperti alcuni decenni fa a Jebel Irhoud, Marocco: il cranio e altre ossa attribuiti a *H.sapiens* arcaico, datati accuratamente a 315.000 *ya*! 100.000 anni prima dei più antichi fossili emersi in Etiopia (vedi sopra). Sono reperti di una specie arcaica o parallela, ma accostabili a *H.sapiens*, con tratti ora moderni, ora più antichi: i fossili di J. Irhoud potrebbero dimostrare che la nostra evoluzione è stata più lunga, diversificata e complessa, dunque parzialmente da riscrivere³⁸.

Percorsi geografici di *H.sapiens*: migrazioni e colonizzazione della terra. Una storia da ricostruire mediante la convergenza dei dati archeologici, paleontologici, genetici e linguistici. Le ricerche degli ultimi 10 anni hanno confermato e rafforzato i dati inerenti tempi, modalità e divaricazioni (o mescolamento) di popolazioni che hanno caratterizzato la diffusione della nostra specie dall'Africa nel mondo, interpretando le indicazioni degli mtAplogruppi³⁹ e del tasso di mutazione genica (Vedi Atti Accademia Lancisiana, 2013). Qui sottolineiamo siti e tempi in cui sono state dimostrate le ibridazioni fra *H.sapiens* e altre specie arcaiche, preesistenti e/o coeve, che hanno determinato nel nostro genoma un'introggressione di geni di cui portiamo ancor oggi, dopo uno sfolgimento inevitabile dovuto al tempo e alle mescolanze, tracce documentate. Fra noi e i Neanderthal gli eventi d'ibridazione iniziarono senz'altro in Medio-Oriente, tra la striscia di Gaza e l'attuale area siro-palestinese, quando furono raggiunte da *H.sapiens* che incontrò i Neanderthal lì presenti da tempo, a partire da 80/85.000 *ya*. La diffusione dell'uomo

moderno verso nord e ovest e ulteriori scambi genetici avvennero nei millenni successivi, fra 47 e 37.000 *ya*, in Russia, Cechia, Balcani e Italia: le nuove introggressioni cancellarono e coprirono in parte le precedenti. L'estinzione dei Neanderthal, com'è noto, avvenne negli ultimi siti ove si erano ritirati, nella penisola iberica, intorno a 33/35.000 *ya*.

Tra il misterioso "Uomo di Denisova", la cui identificazione genetica è stato uno dei maggiori successi scientifici di S. Pääbo, e gruppi umani migrati a est, gli incroci sarebbero avvenuti circa 45.000 *ya* in Siberia: i geni ereditati sono rimasti ancor oggi in diverse popolazioni asiatiche, come i tibetani e i Melanesiani di Papua-Nuova Guinea.

Geni ed evoluzione del linguaggio umano. La ricostruzione del percorso, biologico, cognitivo, relazionale, che, forse, intorno a 50.000 *ya*, ha portato la nostra specie a disporre di una capacità di comunicazione unica ed esclusiva tra gli esseri viventi, rimane lacunosa e più ricca di supposizioni che di dati concreti e dimostrati, e forse rimarrà a lungo, se non per sempre, indefinita⁴⁰.

Durante l'evoluzione, per acquisire questa nuova funzione della nostra specie, una manifestazione cognitiva così diversa e peculiare, è stato indispensabile il concorso di 4 fenomenologie morfo-funzionali:

1. modificazioni anatomiche;
2. coordinazione neuro-muscolare;
3. sviluppo di connessioni e circuiti cerebrali;
4. modificazioni e attivazioni geniche.

Il punto 1 è parzialmente documentabile mediante la comparazione dei fossili tra le varie specie di ominidi, scimmie antropomorfe e altri primati, nonché oggi con la ricostruzione virtuale dei tessuti molli; i punti 2 e 3 rientrano, come detto, nel campo delle congetture e delle ipotesi forse non dimostrabili, mentre il discorso genetico trae nuove acquisizioni dal sequenziamento

dei genomi attuali e antichi. E qui il discorso torna al tanto enfatizzato gene FoxP2, etichettato, impropriamente, il "gene della parola"⁴¹. Situato sul braccio lungo del cromosoma 7, codifica per un fattore di trascrizione importante nello sviluppo di aree cerebrali, del polmone, di strutture neurali connesse con il controllo motorio e l'articolazione del linguaggio. Mutazioni del gene sono associate a dislessia, disprassia e disturbi cognitivi. Presente nella nostra specie, in diversi primati e, analoghi, in altri animali, nel 2007 è stato trovato nel genoma di *H.neanderthalensis* (che forse poteva articolare alcuni suoni), rimettendo in discussione tutte le nostre conoscenze su origine ed evoluzione del linguaggio.

Nella nostra, come in tutte le altre specie, un gene o un gruppo di geni, da soli, non ci fanno parlare o comunicare verbalmente: il linguaggio è una funzione troppo complessa per poter essere regolata geneticamente da un unico elemento, non possiamo accettare l'etichetta semplicistica di "gene della parola". Il linguaggio ha richiesto, per manifestarsi, ben altro. Non esiste un unico gene, umano o ereditato da altre specie del passato, né si sono verificate mutazioni o interconnessioni fra circuiti cerebrali che, da soli ed esclusivi, spieghino il linguaggio. La questione dell'origine e dell'evoluzione del linguaggio umano deve essere affrontata multidisciplinarmente da neurobiologi, linguisti, genetisti, esperti del comportamento umano e animale: dopo aver indagato per anni con una diversificata gamma di approcci metodologici, sembra oggi che non si riesca a trovare nulla di veramente unico in noi stessi, né nel genoma, né nel cervello, che spieghi il linguaggio. Certo è, comunque, che il linguaggio non è direttamente insito nel nostro DNA; abbiamo cercato tracce di qualche evento biologico critico da cui possa essere scaturito, circa 50.000 *ya*, ma tutto è stato vano. Non sono stati trovati tratti distintivi di natura fisiologica, neurologica o genetica, tali da spiegare l'unicità del linguaggio, il quale

sembra emergere da un insieme di capacità, alcune delle quali, singolarmente, condivise con altri animali e, forse, specie ominidi del passato, sia pur in modi e forme differenti. Un'ipotesi unificante è che la complessità del linguaggio umano sia nata da quell'insieme di tratti che caratterizzano la nostra specie, e che chiamiamo CULTURA⁴²⁻⁴⁵.

Mutazioni, evoluzione umana, rimescolamento di popolazioni. La genealogia delle mutazioni umane aiuta ad osservare l'evoluzione lungo centinaia di generazioni, ma può oscurare le complesse interazioni tra ambiente, sopravvivenza e fecondità che si sono manifestate nel passato. Talora sul lungo termine vediamo solo i geni vantaggiosi, come la (persistenza) della lattasi, ma potremmo perdere le dinamiche nel tempo; in altri casi, nuove mutazioni adattative potrebbero essersi diffuse nelle popolazioni, senza espellere le versioni più vecchie dei geni: pertanto, le varietà ancestrali sono rimaste con noi: tutto questo aiuta a ricostruire la storia di ciascun gene. Oggi milioni di persone si spostano da un paese all'altro, portando il tasso di scambio e rimescolamento genetico ad un livello mai verificatosi prima. Saremo nel prossimo futuro un'unica popolazione mondiale variegata, di meticci e caratteri misti? Ognuno dei nostri discendenti sarà un mosaico vivente di storia umana.

Esempi classici di geni i cui alberi genealogici e storia delle mutazioni aiutano a ricostruire l'evoluzione umana, sono: AMILASI SALIVARE, LATTASI, DARC gene e MALARIA, geni per la PIGMENTAZIONE DELLA PELLE, ADATTAMENTO IN QUOTA E RISPOSTA IMMUNITARIA AI PATOGENI.

Il gene della lattasi e le sue mutazioni: un esempio di coevoluzione e intrecci biologici, culturali e sociali. "In passato l'innovazione culturale ha spesso diretto l'evoluzione biologica; la coevoluzione biologica e culturale ha permesso all'uomo di predominare anche su specie più forti fisicamente; ciò fu possibile sia grazie alla

maggior intelligenza che all'accumulo di conoscenze, trasmissibile fra le generazioni, che chiamiamo "cultura" (L.L. Cavalli-Sforza). La persistenza dell'enzima lattasi, di conseguenza la capacità di digerire il latte anche da adulti, è un carattere che dal medio-oriente si è diffuso in Europa e in tutto il mondo, con modalità e tempi non ancora del tutto chiariti, così come le pressioni selettive che l'hanno determinata⁴⁶. Le ipotesi finora più accreditate, almeno per quanto riguarda l'acquisizione di questo tratto genetico in Europa, sono due: diffusione al seguito dell'allevamento, specie di bovini e capro-ovini, dalla rivoluzione neolitica in poi, o, circa 5000 *ya*, portato dai nomadi delle steppe asiatiche. Tra le pressioni selettive che hanno favorito questo carattere, cruciale è stato il ruolo del latte stesso (e dei suoi derivati), ausilio a combattere i patogeni, che andavano diffondendosi nei centri urbani, nuovo modello di convivenza, sempre più grandi, affollati e scarsamente igienici, anche modificando il microbiota intestinale (oggi è dimostrata la sua diversificazione tra i pochi popoli dell'Africa australe ancora nomadi e cacciatori-raccoglitori, le popolazioni di allevatori-agricoltori e le genti urbanizzate): una catena, dunque, che porta a modificare il rapporto tra sistema immunitario, microbiota e alimentazione, condizionando i percorsi di difesa verso vecchi e nuovi patogeni.

Genetica della risposta immunitaria. La materia da sola meriterebbe una trattazione esclusiva ed esaustiva: sintetizziamo qui i punti su cui la ricerca immunogenetica nei prossimi anni si soffermerà maggiormente.

- Differenze di genere nelle malattie infettive, neoplastiche e autoimmuni⁴⁷.
- Rapporti e interferenze tra alimentazione, malattie dismetaboliche, risposta infiammatoria, patogenesi di processi sequenziali quali l'aterosclerosi, geni coinvolti e/o predisponenti.
- Microbioma intestinale: differenze qualitative nelle popolazioni in rapporto alla dieta^{48, 49}, all'ambiente e ad

abitudini di vita, interferenze con la risposta infiammatoria locale e sistemica, rapporto bidirezionale con l'assetto genetico^{50, 51}.

- Adattamenti del sistema immunitario, geneticamente guidati, alla variabilità dell'ambiente microbiologico e alla tipologia di patogeni⁵².
- Ruolo di nuove varianti e mutazioni nella risposta umorale e cellulare⁵³.
- Risposta immune individuale in corso di patologie infettive, neoplastiche e autoimmuni, allestimento di immunoterapie personalizzate o per gruppi di pazienti altamente omogenei: **MEDICINA DI PRECISIONE**.
- Studi in vitro su organoidi, ad es. relazioni fra geni, farmaci, previsione terapeutica.

Infine, in prospettiva, due sono le tipologie d'indagine in cui nuove conoscenze di base e tecnologie innovative avranno ricadute in clinica interessanti e risolutive. **EDITING GENETICO**: come modificare con precisione singoli nucleotidi, o brevi sequenze, finalizzato a **NUOVE TERAPIE GENICHE**. **EPIGENETICA**: non è più sufficiente conoscere solo struttura e localizzazione dei geni, ma anche il grado di espressione, nonché le interferenze esogene, nell'ambito della cosiddetta epigenetica, di cui sono in allestimento per il genoma umano mappe di metilazione e atlanti di attivazione/repressione genica. L'interesse per l'epigenetica è legato al suo ruolo trasversale praticamente in ogni sito del genoma^{54, 55}.

Geni umani, di altri primati e ominidi estinti

H.sapiens condivide con i primati fino al 90 % del DNA; il 10 % divergente riguarda geni, la cui espressione funzionale è implicata, direttamente o su gruppi di altri geni, in meccanismi di regolazione-trascrizione-traduzione, nel cervello, nelle attività motorie e in quelle sensoriali. Per oltre il 30 % il genoma umano è più prossimo a quello dei gorilla che degli scimpanzè,

soprattutto nei geni che regolano lo sviluppo cerebrale e l'udito^{56, 57}. Con i primati e le altre specie *Homo* con cui abbiamo convissuto condividiamo le basi genetiche del sistema ABO. Le popolazioni moderne non africane hanno ereditato il 3/4% del genoma da *H.neanderthalensis*, alcune asiatiche, ma soprattutto melanesiane, il 5/6% dai denisoviani. Gli incroci genetici hanno arricchito qualitativamente il nostro sistema immunitario e il numero di antigeni HLA: ad es., HLAB73 deriva dai Denisoviani, HLA11 dai Neanderthaliani asiatici.

Ominidi estinti, sui cui fossili sono state effettuate con successo indagini genetiche e comparazioni con il DNA dell'uomo moderno, dimostrando antiche ibridazioni fra più specie e un'eredità di geni introgressi ancor oggi attivi e presenti nel nostro genoma, sono essenzialmente *H.neanderthalensis* e uomo di Denisova. Sicuramente le ibridazioni avvennero anche con altre specie: alcune non sono ancora state geneticamente studiate (*H.floresiensis*), altre troppo remote (*H.erectus*), altre sconosciute, "ghost", delle quali ipotizziamo lasciti ancestrali di geni di antichissima origine.

Degli ominidi estinti è stato possibile esaminare mtDNA e nDNA antico; i risultati migliori sono stati ottenuti con reperti di un'epoca compresa fra 80.000 e 35.000 *ya*, comunque a tutt'oggi non è mai stato estratto materiale genetico su fossili più antichi di 100.000 *ya*.

H.sapiens ha convissuto in Europa ed Asia con neanderthaliani e denisoviani, fino all'estinzione di questi ultimi, scambiandosi frammenti genetici di cui l'uomo moderno porta ancora l'impronta nel proprio DNA. L'uomo di Denisova è la prima specie, nel genere *Homo*, identificata mediante esami molecolari sul DNA antico.

Eredità neanderthaliana

- Geni che regolano l'espressione della cheratina, soprattutto in popolazioni d'origine caucasica e mongoloide: da essi è derivato un contributo decisivo a nuovi

caratteri morfologici, quale il colore di pelle e capelli per un nostro miglior adattamento in ambienti extra-africani.

- 18 geni, nella regione cromosomica 3p21.31, connessi con l'assorbimento dei raggi UV (OCA2 e BNC2) e i livelli e il metabolismo della Vit. D, presenti oggi fra gli asiatici orientali e i nord-europei.
- Geni implicati in diverse funzioni dell'immunità innata⁵⁸⁻⁶⁰; tra questi, alleli HLA-A*02, A*26/*66, B*07, B*51, C*07:02 e C*16:02, che hanno contribuito a diversificare il sistema HLA umano, fornendo ulteriori molecole per affrontare nuovi patogeni con cui *H.sapiens* è venuto a contatto durante la sua evoluzione e le sue migrazioni. In seguito ai cambiamenti ambientali, al miglioramento delle condizioni igieniche e al minor impatto dei patogeni, alcuni di questi geni hanno subito riconversioni e adattamenti funzionali e sono nel nostro tempo connessi con una maggior suscettibilità a immunopatie oggi incrementate, quali le malattie autoimmuni (LES, cirrosi biliare primitiva, m. di Crohn, diabete tipo II) e allergiche, l'espressione di citochine (IL 6/18) e un cluster di 3 *Toll Like Receptor* (TLR 1-6-10), ma anche, e qui la relazione rimane oscura, con caratteri morfologici (dimensioni della papilla ottica), abitudini e comportamenti (dipendenza dalla nicotina), patologie psichiche (depressione maggiore). Saggiando macrofagi e monociti provenienti da europei e amerindi, che hanno geni neanderthaliani, le cellule reagiscono ai patogeni intracellulari (virus, micobatteri) in modo più protratto nel tempo, meno effimero, anche se talora con un eccesso di effetti collaterali infiammatori: iperproduzione e tempesta citochinica.
- Geni per la crescita e l'espansione cerebrale^{14, 15, 61-63}: UBR4 (neurogenesi) e PHLPP1 (mielinizzazione); lo sviluppo di aree peculiari e il fenomeno della

globalizzazione, specifici di *H.sapiens*, sono avvenuti anche con l'utilizzo e la riconversione di fattori di regolazione ereditati dai nostri predecessori, oltre che, ovviamente, del tutto nuovi, unicamente umani. Sono poi stati individuati 87 geni implicati nello sviluppo e nelle funzioni cerebrali, che differiscono profondamente fra noi e le altre specie: le differenze non consistono solo nella struttura dei geni e nelle proteine prodotte, ma anche dal grado di espressione (attivazione o repressione) dei geni stessi e dalle influenze epigenetiche, aspetti che solo di recente si è iniziato ad indagare.

- Geni che favoriscono l'assorbimento del complesso vitaminico B.
- Una variante allelica neanderthaliana ereditata dal 30% delle donne europee attuali, codificante per un recettore del progesterone⁶⁴; le portatrici di questa variante producono una maggior quantità di recettore, diversi effetti dell'ormone in gravidanza risultano esaltati e più alto è l'indice di fertilità.

Ricordiamo infine come uno studio recente (*Ellinghaus et al.* 2020) abbia identificato un *cluster* genico sul cromosoma 3 associato ad un alto rischio di danno e insufficienza respiratoria in COVID-19⁶⁵. L'indagine è stata condotta su 3200 pazienti ospedalizzati con malattia di media/severa gravità di varie etnie: interessati sono Eurasiatici e Amerindi, che hanno ereditato il *cluster* dopo antiche ibridazioni fra noi e i Neanderthal (Zeeberg e Pääbo), rispetto agli Africani in cui è quasi assente⁶⁶. Il *cluster* comprende molti geni che codificano per fattori e mediatori dell'immunità innata e della flogosi e potrebbe contribuire alla eccessiva reazione infiammatoria caratterizzata dalla "cytokine storm". Le popolazioni dell'Africa centrale e australe, non interessate da questa eredità, hanno sviluppato una malattia assai meno grave e senza le ricadute dannose di una devastante reazione immuno-infiammatoria.

Eredità denisoviana

- Geni del sistema HLA, diversi da quelli dei Neanderthal, connessi con la presentazione degli antigeni e la successiva risposta a microrganismi patogeni: HLA-B*73, HLA-C*15:05, HLA-A(A*02 e A+119) e HLA-C (C*15 e C*12:02)⁵⁸.
- Un *cluster* di 3 TLR, diverso da quello ereditato dai neanderthaliani.
- Le popolazioni asiatiche e di Papua-Nuova Guinea conservano nel loro genoma un gene, TNFAIP3, che codifica la proteina A20 su induzione di TNF α , nella sola variante allelica di derivazione Denisoviana: contribuisce alla regolazione di diversi aspetti funzionali dell'immunità, quali la tolleranza e/o l'eliminazione di svariati patogeni nell'ambito dei meccanismi di resistenza naturale e dell'azione di citochine pro-infiammatorie. L'allele denisoviano, A20 T108A;I207L, è correlato ad un'esaltazione dei meccanismi dell'immunità naturale senza una concomitante esagerata e dannosa risposta infiammatoria⁶⁷.
- I tibetani attuali hanno ricevuto dai denisoviani una variante genica (EPAS1)⁶⁸ che permette di tollerare la rarefazione dell'ossigeno ad alta quota: regola la concentrazione di emoglobina, impedendo che, in condizioni di prolungata ipossia da altitudine, la poliglobulia non superi determinati valori critici, così come l'ematocrito e la viscosità del sangue e i rischi di trombosi siano contenuti.
- 2 geni, TBX15 e WARS2 sono collegati, in popoli peri-artici, all'adattamento al freddo e alle restrizioni metaboliche di una dieta povera⁶⁹.

Geni esogeni o "alieni" nel DNA umano. Scambi genici

L'inserimento di frammenti più o meno lunghi di genoma, talora perfettamente funzionanti, di una specie in un'altra è un

evento comune nei procarioti, ma è stato da tempo dimostrato anche negli eucarioti: piante e animali, dagli invertebrati ai più complessi vertebrati; nel genoma di questi ultimi, uomo compreso, vi sono alcune centinaia di geni provenienti da altre specie. La trasmissione orizzontale permette a batteri, funghi e virus di evolvere ed acquisire nuove proprietà fenotipiche, così da adattarsi a mutate condizioni ambientali: è un percorso di variabilità genetica che ha contribuito (e contribuisce tuttora) all'evoluzione degli organismi, sì da rivedere le teorie riguardo la speciazione. Questo meccanismo di passaggio e scambio orizzontale di geni, assieme alla trasmissione di informazioni genetiche mediante plasmidi, è implicato anche nella rapida acquisizione di resistenze agli antibiotici da parte dei batteri. Nell'uomo e nei vertebrati questi geni "esogeni" provengono da batteri, virus, protisti e funghi; alcuni codificano proteine enzimatiche coinvolte in reazioni cellulari e in complessi percorsi metabolici. Il nostro genoma contiene DNA o RNA (talora retroscritto a DNA) virale, lasciato di infezioni lontane o più recenti da parte di virus antichi o ancora attivi; si ritiene che questo genoma "alieno" rappresenti almeno l'8% di tutto quello umano⁷⁰. La maggioranza di queste sequenze virali non è più attiva, avendo subito nel tempo mutazioni e ricombinazioni che hanno silenziato moltissimi geni; altri geni, però, hanno conservato la capacità di codificare proteine: mutazioni dei primi e alterazioni delle molecole proteiche sono implicati nell'eziopatogenesi di malattie, quali tumori, disordini autoimmuni, cardiovascolari e miopatie.

Alcuni esempi di interazioni geni virali/DNA umano. (HERVs, Human Endogenous Retroviruses)

- Oncogeni e geni oncosoppressori⁷¹.
- TRIM5, appartenente ad una famiglia di più geni che impediscono la replicazione dei retrovirus nelle cellule

dell'ospite. Fa eccezione HIV, verso il quale la nostra specie non è ancora stata in grado di adottare le idonee e specifiche armi molecolari di inibizione nell'ambito del sistema TRIM, o altri.

- Attraverso modalità di riassetto di sequenze geniche nc, anche molto antiche e d'origine retrovirale, si sono evoluti geni importanti per lo sviluppo del cervello umano, assenti nei primati e negli ominidi estinti di cui abbiamo potuto sequenziare aDNA. È stato inoltre dimostrato che alcuni di questi nc-geni, dopo mutazioni, sono coinvolti nella comparsa e nella progressione di tumori, quali il glioblastoma⁷².
- Sequenze retrovirali inserite anticamente, cruciali per lo sviluppo di cellule embrionali: tra queste, 23, chiamate "linc-RNA o Long Intergenic non coding-RNA", che manifestano un'attività essenziale nella regolazione di molti geni nelle cellule staminali embrionali umane⁷³.
- IRGM, gene antico, presente nel lignaggio dell'uomo e delle grandi scimmie antropomorfe, scomparso e riapparso più volte nei genomi ancestrali⁷⁴. IRGM è uno dei 2 geni tronchi nel genoma umano: codificano entrambi per la GTPasi immuno-correlata, una molecola enzimatica implicata nell'immunità innata verso patogeni intracellulari, come Micobatteri e Salmonelle. È stata individuata una correlazione fra mutazioni del gene e suscettibilità a contrarre il morbo di Crohn.

Significato e interpretazione. Si stima che le sequenze geniche virali, inserite nel nostro genoma, potenzialmente in grado di codificare proteine, potrebbero essere circa un migliaio: perlopiù sono inattive da tempo, o repressi ir/reversibilmente, ma alcune funzionanti soprattutto nell'ambito della risposta immunitaria; in alcuni casi geni (retro)virali silenti, si riattivano in seguito a

modifiche conformazionali in siti cruciali degli stessi geni, stress ambientali, ulteriori infezioni intercorrenti. È innegabile il fatto che i genomi di tutte le specie siano stati fortemente plasmati e condizionati, nella loro evoluzione, dall'impatto e dall'interazione con microrganismi patogeni, virus in particolare, da un punto di vista molecolare e immunitario.

Il nostro genoma è un mosaico di geni antichi e relativamente nuovi, mutati o più o meno conservati, diversi di origine esogena: ospitiamo geni che abbiamo ereditato da nostri antenati ominidi ormai estinti da lungo tempo, o relitti di ancestrali infezioni da parte di microrganismi scomparsi o attuali. Il fatto che questi geni siano sopravvissuti fino a oggi, ne giustifica in ogni caso un ruolo nella nostra fisiologia, o sbagliamo ragionando sempre in chiave finalistica e funzionale? La loro presenza ripropone profonde questioni scientifiche e filosofiche sul concetto di specie. Non sappiamo quanto della loro longevità dipenda da fattori insiti nella loro stessa natura, o da un eventuale intervento funzionale o da fattori ecologici estrinseci; da questi geni possiamo trarre insegnamenti sul ruolo delle contingenze stocastiche nella storia evolutiva degli organismi viventi attuali, noi compresi.

Nel 1942 E. Mayr superò innegabili limiti morfo-funzionali per formulare un nuovo concetto di specie, in base alla possibilità o meno per gli esseri viventi di incrociarsi e fertilizzarsi con successo⁷⁵. Isolamento riproduttivo, infertilità fra gameti di specie diverse, varietà di comportamenti in tema di riproduzione, segnano una linea netta di demarcazione e incompatibilità che distingue le molteplici specie. A lungo, scambi di informazione genetica fra specie diverse sono stati negati. Prima di Darwin, quando il magistero mendeliano era pressoché sconosciuto e i principi della genetica idee nebulse, gli scienziati ritenevano che ogni specie fosse un tipo essenziale, una forma quasi platonica ideale, immutabile: il

naturalista inglese dimostrò invece che le specie cambiano nel tempo e manifestano una considerevole misura di variazione intrinseca, che oggi includiamo nel concetto di MUTAZIONE, ribadendo che tutto questo è un vantaggio, dato che costituisce un substrato essenziale per l'evoluzione. Non poteva sapere, e non lo abbiamo saputo a lungo, che gli esseri viventi, inclusi i vertebrati, sono "aperti" anche ad apporti estrinseci, che oggi identifichiamo nei geni esogeni, ulteriore fonte di VARIABILITÀ'.

Ecco dunque che l'individuazione nei genomi attuali, uomo incluso, di relitti genici esogeni da ominidi estinti o microrganismi quali batteri e virus, mette sia pur parzialmente in discussione i principi di *Mayr*, non tanto nell'impostazione generale, quanto in alcuni dettagli formulati 80 anni fa, senza le attuali conoscenze di genetica. Oggi sappiamo che la distinzione fra specie, soprattutto se distanti nel tempo e nello spazio, è variegata, anatomica, comportamentale, ecologica, biogeografica, sessuale e riproduttiva, evidentemente genetica e molecolare.

Conclusione

Il primo sequenziamento completo di un genoma umano e i dati che abbiamo presentato, frutto di 20 anni di ricerche, sono un punto fermo di partenza per prossimi studi: si aprono per il futuro sfide complesse, ma oltremodo affascinanti. Il genoma umano è fortemente dinamico, non statico, come dimostrano MUTAZIONI e VARIABILITÀ, è anche un'enciclopedia aperta a continui inserimenti, con una frequenza che non è diversa dalle altre specie, semmai persino più diversificata. Attualissimo appare l'approfondimento delle complesse interazioni con l'ambiente, anche in tema di fecondità umana, nonché il contributo dell'epigenetica nell'eziologia dei processi morbosi. Fondamentali sono i progressi tecnologici, dalle macchine sequenziatrici all'apporto informatico (algoritmi, software, A.I., *mechanism learning*), all'*editing genetico*

(CRISPR), alle applicazioni cliniche mediante le nuovissime terapie geniche. La miglior conoscenza del DNA umano e della sua VARIABILITÀ non è utile solo per comprendere meglio l'eziopatogenesi delle malattie, la diversificazione fra patologie congenite e acquisite, la risposta terapeutica soggettiva, ma diverrà la base insostituibile per la medicina di domani, la MEDICINA PERSONALIZZATA.

Da un punto di vista dottrinale e teorico, infine, possiamo oggi ribadire l'UNICITÀ del genoma umano, senza discriminazioni in razze o sottogruppi su base biologica e genetica, pur nella diffusa VARIABILITÀ individuale: UNICITÀ E VARIABILITÀ che hanno costituito le armi vincenti della nostra specie e che devono essere salvaguardate in un ambiente che noi stessi contribuiamo a MUTARE e stravolgere.

Un genoma, dunque, soggetto a cambiamenti, polimorfismi, in una parola MUTAZIONI: chi meglio di altri potrebbe simboleggiare questa realtà biologica se non un'antica divinità etrusca e poi romana, il dio Vertumno? Ovvero la personificazione del *mutare*, l'incontenibile impulso delle cose e delle persone a divenire *altro* da quel che sono. Nasce come divinità agreste etrusca, e poi romana dei cambiamenti delle stagioni e la maturazione dei frutti, la pratica dell'innesto e delle ibridazioni (vera e propria «trasformazione» di una specie vegetale in un'altra, cambiamento biologico, oggi diremmo GENETICO)^{76, 77} e quella dello «scambiare» merci, il mutamento degli eventi e l'atto del cambiare idea. Vertumno si distingue per una sua ineguagliabile virtù: la *metamorfosi*, e come tale è citato da Ovidio e Apuleio; in età classica, il termine impiegato è MUTAZIONE.



Il dio Vertumno. Garden of Anglesey Abbey

BIBLIOGRAFIA

1. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860-921.
2. Nurk S, Koren S, Rhie A et al. The complete sequence of a human genome. *Science* 2022; 376: 44-53.
3. Gonen N, Futtner CR, Lovell-Badge R et al. Sex reversal following deletion of a single distal enhancer of *Sox9*. *Science* 2018; 360: 1469-73.
4. Walsch K, Sano S, Horitani K et al. Hematopoietic loss of Y chromosome leads to cardiac fibrosis and heart failure mortality. *Science* 2022; 377: 292-7.
5. Ezkurdia I, Juan D, Rodriguez JM et al. Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Hum Mol Genet* 2014; 23: 5866-78.
6. Fox D. Junk DNA deforms salamander bodies. *Scientific American* 2022; 326: 2.
7. Levi-Strauss C. *Razza e storia e altri studi di antropologia*. Torino: Einaudi, 1968.
8. Cavalli-Sforza LL, Menozzi P, Piazza A. *Storia e geografia dei geni umani*. Milano: Adelphi, 1997.
9. Levi-Montalcini R, Tripodi G. *Eva era africana*. Roma: Gallucci, 2005.
10. Orphanet Italia. Registro Nazionale Malattie rare. Report 2016. www.orpha.net.
11. Nowakowski TJ, Pollen AA, Kriegstein AR et al. Molecular identity of human "Outer Radial Glia" during cortical development. *Cell* 2015; 163: 55-67.
12. Ataman B, Boulting GL, Greenberg ME et al. Evolution of Osteocrin as an activity-regulated factor in the primate brain. *Nature* 2016; 539: 242-7.
13. Sousa AMM, Zhu Y, Raghanti MA et al. Molecular and cellular reorganization of neural circuits in the human lineage. *Science* 2017; 358: 1027-32.
14. Fiddes T, Lodewijk GA, Haussler D et al. Human-specific *NOTCH2NL* genes affect Notch signaling and cortical neurogenesis. *Cell* 2018; 173: 1356-69.
15. Suzuki IK, Gacquer D, Vanderhaegen P et al. Human-specific *NOTCH2NL* genes expand cortical neurogenesis through Delta/Notch regulation. *Cell* 2018; 173: 1370-84.
16. Zwir I, Del Val C, Hintsanen M et al. Evolution of genetic networks for human creativity. *Mol Psy* 2022; 27: 354-76.
17. Tattersall I. *Becoming human: evolution and human uniqueness*. New York: Harcourt Brace, 1998.
18. Lalli MA. Origin of the PSEN1 E280E mutation causing early-onset Alzheimer's disease. *Alzheimer's Dementia* 2014; 10: 277-83.
19. Tarawneh R, Head D, Fagan A et al. Cerebrospinal fluid markers of neurodegeneration and rates of brain atrophy in early Alzheimer disease. *JAMA Neurol* 2015; 72: 656-65.
20. Suarez-Calvet M, Araque Caballero MA, Kleinberger G et al. Early changes in CSF sTREM2 in dominantly inherited Alzheimer's disease occur after amyloid deposition and neuronal injury. *Sc Transl Med* 2016; 8: 369.
21. Hoch NC, Hanzlikova H, Rulten SL et al. *XRCC1* mutation is associated with PARP1 hyperactivation and cerebellar ataxia. *Nature* 2016; 541: 87-91.
22. Belli F. Biomarcatori nella diagnosi precoce e nel monitoraggio del morbo di Alzheimer: utopia o prossima realtà? *Policlinico Sez Med* 2021; 128: 71-82.
23. Novarino G, El Fishawy P, Kayserili H et al. Mutations in BCKD-kinase lead to a potentially treatable form of autism with epilepsy. *Science* 2012; 338: 394-7.
24. Belli F. La medicina che verrà. La rivoluzione delle biotecnologie in un percorso clinico personalizzato. *Atti Acc Lancisiana* 2020; 64: 133-60.
25. Sekar A, Bialas AR, de Rivera H et al. Schizophrenia risk from complex variation of complement component 4. *Nature* 2016; 530: 177-83.
26. Van Os J. The environment and schizophrenia. *Nature* 2010; 468: 203-12.
27. Gierman HJ, Fortney K, Roach JC et al. Whole-Genome sequencing of the world's oldest people. *PLoS ONE* 2014; 9: e112430.
28. Khan SS, Shah SJ, Vaughan D et al. A null mutation in *SERPINE1* protects against biological aging in humans. *Science Advances* 2017; 3: eaao1617.
29. Filer D, Thompson MA, Takhaviev V. RNA polymerase III limits longevity downstream of TORC1. *Nature* 2017; 552: 263-7.

30. Garagnani P, Marquis J, Delledonne M et al. Whole-Genome sequencing analysis of semi-centenarians. *eLIFE* 2021; 10: e57849.
31. Sulem P, Helgason H, Stefansson K et al. Identification of a large set of rare complete human knockouts. *Nat Genet* 2015; 47: 448-52.
32. Gudbjartsson DF, Helgason H, Stefansson K et al. Large-scale Whole-Genome sequencing of the Icelandic population. *Nat Genet* 2015; 47: 435-44.
33. Markram H, Muller E, Ramaswamy S et al. Reconstruction and simulation of neocortical microcircuitry. *Cell* 2015; 163: 456-92.
34. Kathiresan S, Saleheen D, Natarajan P et al. Human knockouts and phenotypic analysis in a cohort with a high rate of consanguinity. *Nature* 2017; 544: 235-9.
35. Ardile K, Stranger BE, Montgomery SB et al. Enhancing GTEx by bridging the gaps between genotype, gene expression, and disease. *Nat Gen* 2017; 49: 1664-70.
36. Belli F. I geni raccontano la nostra storia. *Antropologia molecolare e genetica svelano il percorso di H.sapiens*. Atti Acc Lancisiana 2013; 57: 149-58.
37. Belli F. *Geni, storia, evoluzione*. Roma: Aracne, 2015.
38. Scerri EML. Did our species evolve in subdivided populations across Africa, and why does it matter? *Trends Ecol Evol* 2018; 33: 582-94.
39. Martiniano R, De Sanctis B, Hallast P, Durbin R. Placing aDNA sequences into reference phylogenies. *Mol Biol Evol* 2022; 39: msac017.
40. Bolhuis JJ, Tattersall I, Chomsky N, Berwick RC. Language: UG or not to be, that is the question. *PLoS Biol* 2015; 13: e1002063.
41. Krause J, Lalueza Fox C, Orlando L et al. The derived FOXP2 variant of modern humans was shared with Neandertal. *Curr Biol* 2007; 17: 1908-12.
42. Hauser MD, Yang C, Lewontin RC et al. The mystery of language evolution. *Fron Psych* 2014; 5: 1-12.
43. Lieberman P. Language did not spring forth 100,000 years ago. *PLoS Biol* 2015; 13: e1002064.
44. Gomez-Olivencia A, Barash A, Garcia-Martinez D et al. 3D virtual reconstruction of the Kebara 2 Neandertal thorax. *Nat Comm* 2018; 9: art.4387.
45. Lyn H. The question of capacity: why enculturated and trained animals have much to tell us about the evolution of language. *Psyc Bull Rev* 2017; 24: 85-90.
46. Burger J, Link V, Blöcker J et al. Low prevalence of lactase persistence in Bronze Age Europe indicates ongoing strong selection over the last 3000 years. *Curr Biol* 2020; 30: 4307-15.
47. Qu K, Zaba LC, Giresi PG et al. Individuality and variation of personal regulomes in primary human T cells. *Cell Syst* 2015; 1: 51-61.
48. Johnson RJ, Andrews P. Fructose, uricase, and the back to Africa hypothesis. *Evol Anthr* 2010; 19: 250-7.
49. Kratzer JT. Evolutionary history and metabolic insights of ancient mammalian uricases. *PNAS* 2014; 111: 3763-8.
50. Ruiz-Calderon JF, Cavallin H, Song A et al. Walls talk: microbial biogeography of homes spanning urbanization. *Sci Adv* 2016; 2: e1501061.
51. Dominguez-Bello M, Knight R, Gilbert JA et al. Preserving microbial diversity. *Science* 2018; 362: 33-4.
52. Netea MG, Domínguez-Andrés J, Kuipers Y et al. Evolution of cytokine production capacity in ancient and modern European populations. *eLife* 2021; 10: e64971.
53. Zhang FR, Huang W, Ghen SM et al. Genomewide association study of leprosy. *N Engl J Med* 2009; 36: 2609-18.
54. Skinner MK, Manikkam M, Guerrero-Bosagna C. Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends End Met* 2010; 21: 214-22.
55. Daxinger L, Whitelaw E. Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat Rev Gen* 2012; 13: 153-62.
56. Lomax-Boyd J, Skove SL, Rouanet JP et al. Human-Chimpanzee differences in a *FZD8* Enhancer alter cell-cycle dynamics in the developing neocortex. *Curr Biol* 2015; 25: 772-9.
57. Florio M, Albert M, Taverna E et al. Human-specific gene *ARHGAP11B* promotes basal progenitor amplification and neocortex expansion. *Science* 2015; 347: 1465-70.
58. Dannemann M, Andres AM, Kelso J. Introgression of Neandertal and Denisovan like haplotypes contributes to adaptive

- variation in human Toll Like Receptors. *Am J Hum Gen* 2016; 98: 22-33.
59. Deschamps M, Laval G, Fagny M et al. Genomic signatures of selective pressures and introgression from archaic hominins at human innate immunity genes. *Am J Hum Gen* 2016; 98: 5-21.
 60. Quach H, Rotival M, Barreiro L et al. Genetic adaptation and Neandertal admixture shaped the immune system of human populations. *Cell* 2016; 167: 643-56.
 61. Gunz P, Tilot AK, Wittfeld K et al. Neandertal introgression sheds light on modern human endocranial globularity. *Curr Biol* 2018; 29: 120-7.
 62. Florio M, Albert M, Taverna E et al. Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science*; 2015: 347: 1465-70.
 63. Muotri AR, Trujillo CA, Rice EK et al. Reintroduction of the archaic variant of *NOVA1* in cortical organoids alters neurodevelopment. *Science* 2021; 371: eaax2537.
 64. Zeberg H, Kelso J, Pääbo S. The Neanderthal progesterone receptor. *Mol Biol Evol* 2020; 37:2655-60.
 65. Severe Covid-19 GWAS Group, Ellinghaus D, Degenhardt F, et al. Genome-wide association study of severe Covid-19 with respiratory failure. *NEJM* 2020; 383: 1522-34.
 66. Zeberg H, Pääbo S. The major genetic risk factor for severe COVID-19 is inherited from Neandertals. *Nature* 2020; 587: 610-2.
 67. Zammit NW, Siggs OM, Gray PE et al. Denisovan, modern human and mouse *TNFAIP3* alleles tune A20 phosphorylation and immunity. *Nat Immun* 2019; 20: 1299-310.
 68. Huerta-Sanchez E, Jin X, Nielsen R et al. Altitude adaptation in Tibetans caused by introgression of Denisovian-like DNA. *Nature* 2014; 512: 194-7.
 69. Racimo F, Gokhman D, Fumagalli M et al. Archaic adaptive introgression in *TBX15/WARS2*. *Mol Biol Evol* 2017; 34: 509-24.
 70. Belli F. Geni “esogeni” nel DNA umano. Origine, significato e impatto clinico. *Policlinico Sez Med* 2020; 127: 47-59.
 71. Alcazer V, Bonaventura P, Depil S. Human endogenous retroviruses (HERVs): shaping the innate immune response in cancer. *Cancers (Basel)* 2020; 12: 610.
 72. Frenster JD, Inocencio J, Placantonakis DG. Lentiviral transduction of primary human glioblastoma cultures. *Methods Mol Biol* 2018; 1741:81-9.
 73. Durruthy J, Sebastiano V, Wossidio M et al. The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming. *Nat Gen* 2016; 48: 44-52.
 74. Bekpen C, Marques-Bonet T, Eichler E et al. Death and resurrection of the IRGM gene. *PLoS Genet* 2009; e1000403.
 75. Mayr E. Systematics and the origin of species from the viewpoint of a zoologist. New York: Columbia UP, 1942.
 76. Ovidio PN. Le metamorfosi, a cura di V.Sermonti. Milano: Rizzoli, 2014.
 77. Apuleio L. Le metamorfosi o l’Asino d’oro. Milano: Rizzoli, 2005.
- Prof. Francesco Belli, Accademico; già Dirigente Medico Microbiologia e Virologia A.O. San Camillo-Forlanini, Roma; già Docente Immunologia C.d.L. Biotecnologie, “Sapienza” Università di Roma
- Per la corrispondenza: f.belli11@virgilio.it