

IL PROGETTO GENOMA UMANO

Indice

1	Che cos'è il genoma?	2
1.1	Il sequenziamento	3
1.1.1	La tecnica del sequenziamento	4
1.2	Le unità di misura	5
2	Le finalità del PGU	5
3	La storia del Progetto	6
4	L'organizzazione - HUGO	7
5	I progetti ancillari	8
5.1	Variabilità genetica di popolazione	8
5.2	Il sequenziamento di altri genomi	9
6	Dove si trovano i risultati?	9
6.1	La gene bank	9
6.2	I siti web	10
7	Questioni bioetiche e aspetti commerciali	10
7.1	A chi appartiene il genoma umano?	10
7.2	Un ritorno dell'eugenetica?	11
7.3	Il parere del CNB italiano	11

Introduzione

Le finalità principali del progetto riguardano l'acquisizione di informazioni utili per individuare l'eventuale implicazione di alterazioni nella sequenza del DNA nello sviluppo di patologie genetiche nell'uomo, e per comprendere le basi genetiche dell'evoluzione e del funzionamento dell'organismo umano



Il Progetto Genoma Umano (The Human Genome Project) è un progetto internazionale il cui scopo finale è la descrizione completa del *genoma* umano mediante il sequenziamento, cioè mediante l'identificazione della disposizione delle lettere del *codice genetico*, le basi nucleotidiche, lungo tutta la doppia elica del *DNA*. Nel progetto sono convogliati sia gli sforzi della ricerca pubblica di molte nazioni (il cui fine è unicamente quello di ampliare le conoscenze per porre le basi per ricerche future, e che operano quindi coordinando il lavoro tra i diversi laboratori) sia quelli di aziende private che operano in maniera centralizzata.

1 Che cos'è il genoma?

Il genoma è l'insieme di tutte le informazioni genetiche depositate nella sequenza del DNA contenuto nel *nucleo* delle cellule sotto forma di *cromosomi*.

Ogni cromosoma è costituito da un lungo filamento di DNA organizzato in una complessa struttura tridimensionale.

Nel genoma risiedono sia sequenze che non sembrano avere apparentemente una funzione, il cosiddetto “junk DNA” (DNA spazzatura), come le sequenze fossili che si sono inserite nel nostro genoma milioni di anni fa ma senza assolvere alcun compito, sia geni, cioè le sequenze in cui risiede l'informazione genetica per la sintesi delle proteine, che stabiliscono le caratteristiche peculiari di ogni individuo.

Si stima che nel genoma umano siano presenti circa 50000 geni.

La sequenza del DNA risulta dalla disposizione lineare di quattro molecole differenti, i nucleotidi o basi.

I nucleotidi vengono denominati con i nomi adenina, timina, citosina, guanina e vengono indicati rispettivamente con le lettere maiuscole A, T, C, G.

Un esempio di sequenza potrebbe pertanto essere ATC-TCC-CAG-TTA-GCT.

A ogni tripletta di nucleotidi corrisponde una determinata informazione per la sintesi delle proteine, i macchinari che eseguono tutte le funzioni fondamentali delle cellule.

Se una di queste lettere cambia, subisce cioè una mutazione, l'informazione genetica può risultare alterata e le *proteine* potrebbero

non venire più prodotte in modo funzionale causando lo sviluppo di una malattia genetica.

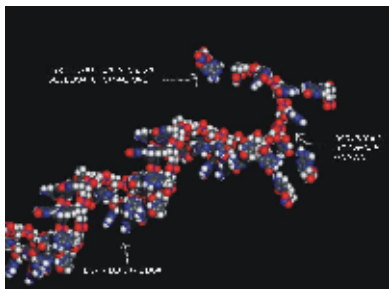
Un esempio di questo fenomeno è la talassemia, causata proprio dalla mutazione in una base del DNA, che conduce alla formazione di un'emoglobina, la molecola che trasporta l'ossigeno, non funzionale.

Il link *Mutazioni puntiformi* permette di scaricare gioco divertente in formato eseguibile che aiuta a comprendere il significato delle mutazioni puntiformi.

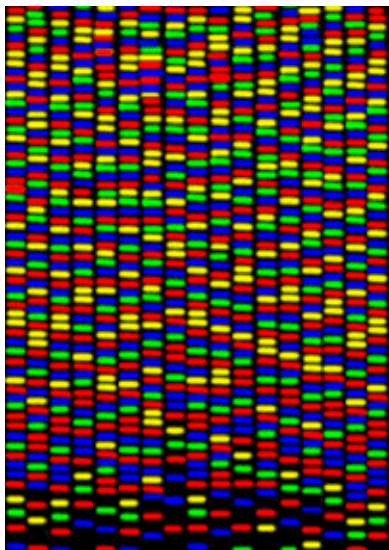
Si stima che l'intero genoma umano sia costituito da oltre 3,12 miliardi di queste basi messe in sequenza, e sequenziarlo vuol dire individuare e ordinare ciascuna di queste lettere.

Il genoma è diverso tra individui, ma solo dell'uno per mille, e questa differenza è sufficiente per riflettersi nella variabilità che si osserva non solo tra le persone, ma anche tra i gruppi etnici.

Questo vuol dire che non esiste un genoma universale da sequenziare, infatti il Progetto Genoma Umano intende sequenziare il genoma di un'unica persona e poi andare a ricercare le differenze tra i vari individui, in particolar modo tra soggetti portatori di determinate malattie e soggetti sani. Questo anche dell'ambito di un ambizioso progetto, quello di ricercare nel DNA la storia dell'evoluzione e delle migrazioni delle popolazioni umane.



1.1 Il sequenziamento



Il Progetto Genoma Umano prevede l'analisi del genoma attraverso la tecnica del sequenziamento del DNA.

Il sequenziamento consiste nell'individuare e ordinare tutti i nucleotidi che costituiscono il nostro patrimonio genetico così come sono posizionati nel genoma.

I nucleotidi sono quattro e sono le molecole di base con le quali è costruito il DNA, sono indicati anche con il termine di basi e sono l'adenina (A), la timina (T), la guanosina (G), la citosina (C).

Il termine mappare si riferisce invece all'individuazione della localizzazione fisica di un gene su un cromosoma.

Sequenziare vuol dire quindi “leggere” l'ordine in cui sono disposte lungo il DNA le basi, cioè le lettere del codice genetico.

Il sequenziamento costituisce una tappa fondamentale per la comprensione del ruolo delle varie parti del genoma umano, è un trampolino di partenza per decodificare il nostro patrimonio genetico.

Infatti il semplice sequenziamento non ci fornisce informazioni direttamente applicabili per conoscere i meccanismi alla base dei processi fisiologici e patologici dell'uomo, ma rappresenta uno strumento grazie al quale sarà più sem-

plice in futuro identificare il ruolo delle diverse porzioni di DNA.

Questo tipo di conoscenze permetterà di identificare le eventuali differenze genetiche tra persone affette da patologie e persone sane, e in futuro l'individuazione di queste divergenze potrebbe essere utile non solo per diagnosticare una malattia prima dell'insorgenza, e pertanto prevenirla dove possibile, ma anche per ideare strategie per curare definitivamente questi soggetti correggendo l'alterazione direttamente a livello del genoma.

Dal punto di vista tecnico per effettuare un sequenziamento è necessario isolare un frammento di DNA, in quanto con le tecnologie ora disponibili non è possibile sequenziare tutto il genoma direttamente.

Quindi si procede spezzettando il genoma in vari frammenti che vengono inseriti in vettori, sequenze di DNA in grado di ospitare frammenti genici esterni e di permetterne l'amplificazione, cioè la produzione in grande numero.

A questo punto, isolato un frammento, si procede con il sequenziamento vero e proprio. I metodi di isolamento dei frammenti di DNA sono diversi, e all'interno del Progetto Genoma Umano nell'ambito della ricerca pubblica, i vari laboratori si sono divisi i vari frammenti di cromosomi da sequenziare.

Gli scienziati della Celera Genomics hanno proceduto lavorando su tutto il genoma utilizzando la tecnica messa a punto per questo scopo da Craig Venter. Questa tecnica è stata denominata “whole genome shotgun sequence technique” e si basa sull'analisi contemporanea dei due filamenti di DNA che compongono la doppia elica, dopo aver ottenuto dei frammenti dall'intero genoma.

I cromosomi umani vengono frammentati in modo casuale in pezzi di 2000 e 10000bp in lunghezza. Quindi i frammenti così ottenuti vengono inseriti in un vettore plasmidico, cioè un anello di DNA che può essere inserito in batteri del tipo *E. coli*, in modo che questi ne producano un'elevata quantità (milioni di copie) necessaria per il sequenziamento.

Infine si procede al sequenziamento che viene effettuato da entrambi i lati dei filamenti di DNA in contemporanea.

La grande mole di dati prodotti con questa tecnica viene infine elaborata da cervelli elettronici molto potenti utilizzando sofisticati programmi di elaborazione dati.

In questo modo ovviamente si ottengono le sequenze di frammenti di DNA di cui non si conosce l'ordine. Molti di questi frammenti, essendo stati generati in maniera casuale, saranno parzialmente sovrapponibili e pertanto sarà possibile, grazie sia ad un lavoro manuale che all'ausilio dei computer, ordinare tutti i frammenti e ottenere un'unica lunga sequenza di DNA per ogni cromosoma, cioè la sequenza completa del genoma umano.

1.1.1 La tecnica del sequenziamento

Isolato il frammento di DNA si procede al sequenziamento vero e proprio.

In questa fase viene utilizzata la DNA polimerasi. La DNA polimerasi è un enzima in grado di sintetizzare un nuovo filamento di DNA durante il processo di duplicazione, processo che permette di ottenere due copie del genoma di una cellula prima che questa si divida per generare due cellule figlie.

Inoltre la DNA polimerasi assolve anche funzioni di riparazione del nuovo filamento sintetizzato in caso di errori di copiatura. Questo enzima è stato scoperto nel 1955 da Arthur Kornberg e collaboratori alla Stanford University.

In particolare la DNA polimerasi è in grado di sintetizzare un nuovo filamento di DNA partendo da un singolo filamento al quale è associato un oligonucleotide, denominato primer.

La DNA polimerasi attacca un nucleotide a questa sequenza copiando in maniera complementare quella originaria, che funge così da stampo. E quindi procede in modo sequenziale fino alla fine del filamento generando una nuova doppia elica.

Nella miscela di reazione utilizzata si in-

seriscono quindi il DNA da sequenziare, la DNA polimerasi, gli oligonucleotidi e i nucleotidi.

Oltre alle basi normali vengono però inserite anche delle basi particolari, i didesossinucleotidi, molecole che la DNA polimerasi può attaccare al filamento nascente del DNA, ma che formano una specie di cappuccio che impedisce l'aggiunta di ulteriori nucleotidi.

Si preparano diverse miscele di reazione contenenti ciascuna diversi nucleotidi didesossi, che verranno incorporati in maniera casuale nel filamento di DNA nascente dalla DNA polimerasi, facendo sì che si generino frammenti di tutte le lunghezze possibili.

Si fanno migrare i prodotti di questa reazione attraverso un gel grazie all'applicazione di una corrente elettrica (il DNA è carico negativamente) in modo che si separino in base alle dimensioni.

I gel che vengono utilizzati presentano infatti dei pori attraverso i quali le molecole più piccole passano più agevolmente e quindi migrano più velocemente (e più distalmente) rispetto alle molecole grandi.

A questo punto si procede alla "lettura" del gel. Nel sequenziamento manuale vengono utilizzate delle sostanze radioattive per cui i gel quando vengono posti a contatto con delle pellicole fotografiche sono in grado di impressionarle.

Il risultato è una lastra che presenta una serie di bande, ciascuna corrispondente ad un frammento di diversa lunghezza.

La lettura viene effettuata secondo la regola che l'ultimo nucleotide del frammento deve essere il complementare della base didesossi presente nella reazione (quindi se nella miscela c'era una didesossiadenina, l'ultimo nucleotide del frammento doveva essere una timina, in quanto l'adenina si associa sempre alla timina).

In questo modo vengono segnate le ultime basi di tutti i frammenti, per ogni miscela di reazione, e queste lette nell'ordine con cui sono migrati i frammenti corrispondenti costituiscono la sequenza.

Il sequenziamento viene effettuato al giorno d'oggi per lo più in maniera automatica, at-

traverso il sequenziatore automatico, e invece delle sostanze radiattive si usano dei nucleotidi colorati.

Il DNA preparato per il sequenziatore automatico viene anch'esso caricato su un gel e fatto migrare, ma la lettura viene fatta direttamente durante la migrazione grazie ad un lettore ottico collegato ad un computer.

Esistono quindi degli appositi programmi che permettono di trasformare le informazioni visive, in termini di densità ottica, in informazioni genetiche, e quindi il computer permette di ottenere direttamente dal gel la sequenza, senza la necessità di un operatore che la legga personalmente.

1.2 Le unità di misura

Esistono più unità di misura utilizzate per quantificare il DNA, la più usata è quella che fa riferimento al numero delle basi e che usa i simboli bp, kb, Mb (per esteso: base pairs -coppia di basi, kilobasi, megabasi).

Per esempio un frammento costituito da 1000000 di basi, o meglio paia di basi essendo il DNA costituito da due eliche in cui i nucleotidi si affacciano a coppie, può essere indicata come 1000000bp, 1000kb, o 1Mb.

Un'altra unità di misura utilizzata è il cM, centiMorgan, che non indica una lunghezza fisica, ma una lunghezza definita genetica. Infatti è un indice della distanza fra due geni in base alla probabilità che negli eventi di scambio che avvengono durante la generazione dei gameti i due geni restino uniti. Più sono vicini e più sarà difficile che vengano separati.

2 Le finalità del PGU

A cosa serve il Progetto Genoma Umano?

L'idea di fondo del Progetto Genoma Umano è l'acquisizione di conoscenze che potranno essere di fondamentale importanza nel comprendere i meccanismi della genetica umana e l'implicazione dei geni nello sviluppo delle malattie

umane. Per perseguire questo fine è necessario isolare dei frammenti di genoma con tecniche più o meno sofisticate, la costruzione di mappe genetiche sempre a più alta risoluzione di questi frammenti fino ad ottenere una serie di frammenti sequenziati che possano essere allineati a descrivere l'intero genoma.

Adesso che il genoma umano è stato sequenziato il prossimo obiettivo, una volta completata la mappatura, è quello di associare ad ogni sequenza genica una funzione, considerando che buona parte del DNA non ha probabilmente una funzione ma rappresenta solo un fossile genetico. Questo lavoro richiederà molto tempo anche perché alcuni geni possono dar luogo a più proteine diverse o ad una proteina unica che però può avere una funzione diversa a seconda del tessuto o dello stadio embrionale in cui viene espressa.

In questo tipo di ricerca rivestono un'importanza fondamentale gli animali, in particolar modo i topi knock-out, cioè animali in cui è stato eliminato artificialmente un gene. In questo modo si può osservare cosa accade in mancanza di determinati geni e si può così dedurre la funzione delle proteine prodotte dai vari geni. E' interessante notare che questi esperimenti possono dare dei risultati interessanti e traslabili alla genetica umana per via dell'elevata somiglianza che esiste fra i geni dell'uomo e quelli del topo e di altri animali da laboratorio. Un'analisi di questo tipo richiede in genere almeno tre anni per ogni singolo gene, quindi è un lavoro molto lungo, a differenza di quello del sequenziamento che è stato fatto prettamente in maniera automatizzata, ma è estremamente interessante perché da dei risultati direttamente applicabili alla medicina umana.

Le informazioni che si potranno ottenere dal sequenziamento genico e dall'individuazione delle funzioni permetteranno di identificare la predisposizione a sviluppare una malattia su base genetica in epoca prenatale o nell'adulto, in fase asintomatica permettendo di prevenirne o ritardarne l'insorgenza. Inoltre potranno essere approfonditi per esplorare e capire la natura genetica dei fenomeni biologici nell'uomo, in particolare quelli patologici, permettendo di

ideare delle terapie efficaci, o dove è possibile, addirittura procedere con la terapia genica. Queste possibilità cambierebbero il modo di vedere la medicina, che dal trattamento per la cura di malattie conclamate passerebbe alla prevenzione su base genetica individuale.

La conoscenza della sequenza del genoma umano è un importante passo verso la comprensione della natura biologica dell'uomo, ma bisogna considerare che il destino di ogni singolo individuo non è dettato solo dalla genetica ma anche dall'ambiente, quindi anche quando saranno note tutte le funzioni delle proteine prodotte a partire dai nostri geni, non sarà possibile prevedere come si svolgerà la vita di ogni singolo individuo, perché questa verrà plasmata dal caso.

3 La storia del Progetto

- Nel 1983 i due più importanti laboratori per lo sviluppo della genetica umana in relazione allo sviluppo delle armi nucleari, a Los Alamos e a Livermore, cominciarono a lavorare al Progetto Biblioteca Genetica. In particolare la DOE (US Department of Energy) era interessata agli effetti delle mutazioni nell'ambito dei programmi nucleari civili e militari

- Nel Dicembre 1984 ad Alta, nello Utah durante un convegno sponsorizzato dalla DOE, il cui scopo era valutare l'efficacia delle tecnologie utilizzate per identificare le mutazioni, si avanzò l'ipotesi che per poter identificare ad alta efficienza le alterazioni nella sequenza del DNA era necessaria l'impostazione di un progetto molto complesso e costoso. Il biologo Robert Sinsheimer propone di fondare un istituto per studiare la sequenza del genoma

- Nel Marzo 1986 la DOE sponsorizza un altro convegno, a Santa Fe, in New Messico, in cui viene evidenziata l'importanza della necessità e della fattibilità di ordinare e sequenziare tutti i frammenti di DNA del genoma umano isolati fino a quel momento, e di continuare a individuarne altri. Dulbecco scrive un articolo su Science in cui riporta la necessità di conoscere la

sequenza dei geni per poter sconfiggere il cancro e capire molti fenomeni biologici.

- 1987 la DOE pubblica Report on the Human Genome Initiative: in questa pubblicazione vengono individuati tre fronti che dovevano essere potenziati:

- la realizzazione di mappe fisiche dei cromosomi umani

- lo sviluppo di tecnologie di supporto per la ricerca sul genoma umano

- il potenziamento della divulgazione delle informazioni, e in particolare la realizzazione di un apposito software di database

- 1988 la US Office of Technology Assessment e il National Research Council pubblicano diversi articoli in merito

- 1988 l'NIH (US National Institutes of Health) istituisce un organo competente per questa materia all'interno della sua organizzazione, l'Office of Human Genome Research (poi rinominato come National Center for Human Genome Research)

- 1988 il Congresso americano approva in via ufficiale il Progetto Genoma Umano della durata di 15 anni con inizio nel 1991. Inizialmente si stimò che sarebbero serviti 15 miliardi di dollari.

- 1999: nel corso dell'anno Celera Genomics, una società privata americana, fondata da Craig Venter impegnata in ricerche di genetica, presenta 6500 richieste di brevetti

- dicembre 1999: su Nature viene pubblicata la sequenza completa del cromosoma 22, uno dei più piccoli fra i cromosomi. A completarne il sequenziamento sono stati gli scienziati del Sanger Center, il centro di ricerca che si stima stia effettuando quasi un terzo di tutto il lavoro del settore pubblico all'interno del Progetto Genoma Umano (The Sanger centre (Gb), <http://www.sanger.ac.uk>).

- 10 gennaio 2000: Greg Venter annuncia che la Celera Genomics ha sequenziato e archiviato le sequenze di diversi frammenti di genoma che, secondo i loro calcoli, dovrebbero coprire il 90% del genoma umano.

- 24 Marzo 2000: su Science esce la sequenza completa del Genoma di *Drosophila Melanogaster*, da parte della Celera Genomics, in collaborazione con gli scienziati del Berkeley *Drosophila* Genoma Project (BDGP). Durante il progetto di sequenziamento hanno identificato 14000 geni.

- Giovedì 6 aprile 2000: (Nature Biotechnology, 18, 475, 2000) annuncio del completamento della sequenziamento del genoma umano da parte della Celera Genomics (<http://www.celera.com>), progetto intrapreso da questa società l'8 settembre 1999. Hanno completato il sequenziamento, e completeranno la mappatura entro l'anno. Il genoma umano si è rivelato composto da 3,12 miliardi di coppie di basi nucleotidiche. Il sequenziamento è stato effettuato sequenziando dei frammenti di DNA ottenuti dal genoma, che sono fra loro sovrapponibili in modo che sarà possibile ordinarli, e le sequenze ottenute ricoprono il genoma 4,6 volte (cioè sono state sequenziate oltre 14 miliardi di basi). La Celera Genomics continuerà il sequenziamento per l'identificazione degli SNP (single nucleotide polymorphisms), cioè di quei singoli nucleotidi che cambiano da un individuo all'altro e che in alcuni casi determinano la suscettibilità a sviluppare malattie.

- 10 aprile 2000: su Nature, annuncio del completamento del sequenziamento al 99,7% del CR21 (rimangono tre piccoli buchi di 100kb), il più piccolo dei cromosomi, la cui presenza in tre copie causa la Sindrome di Down. La parte sequenziata è di 33546361bp, contiene 127 geni noti, 98 sequenze corrispondenti probabilmente ad altrettanti geni, 59 pseudogeni

- 26 Giugno 2000: Conferenza stampa alla Casa Bianca di Clinton con i rappresentanti della Celera Genomics per annunciare il completamento di una bozza della sequenza del genoma in cui sono già riportate le localizzazioni di molti geni sui cromosomi. Il lavoro di assemblamento delle sequenze nucleotidiche è stato effettuato grazie all'utilizzo di computer per un totale di 20000 CPU-ore, ed è stato ricontrollato utilizzando una tecnica diversa dalla shotgun, la tecnica utilizzata per il sequenziamento, denominata "Regional Assembler" che si basa sul

posizionamento delle sequenze contenute negli YAC, i cromosomi artificiali di lievito in cui sono state inserite larghe sequenze di genoma umano, sui cromosomi. In questo modo si ottengono dei frammenti sequenziati denominati "scaffolds" che ricostituiscono i 24 cromosomi umani.

4 L'organizzazione - HUGO

Collaborano al Progetto Genoma Umano circa 50 nazioni. Gli Stati Uniti, che sono i maggiori contribuenti, hanno come supportatori la DOE e l'NIH.

Moltissimi Stati europei hanno individuato la necessità di contribuire a questo progetto, e la stessa CE ha proposto dei progetti transnazionali.

Molte nazioni hanno creato dei propri programmi di ricerca. Le prime nazioni coinvolte sono state il Regno Unito, e la Francia, quindi la Danimarca, la Germania, l'Italia, l'Olanda.

Tra i primi ad appoggiare il progetto va annoverato il Giappone. Partecipano inoltre anche URSS (CIS), Australia, Canada, Korea, Nuova Zelanda.

Per coordinare i lavori di tutte le nazioni è stato fondata nel 1988 l'Human Genome Organization (HUGO), che ha attualmente tre sedi: HUGO Europa (Londra), HUGO America (Bethesda, USA), HUGO Pacifico (Tokyo).

I ricercatori che fanno parte di questo progetto hanno l'obbligo di immettere le sequenze da loro identificate nel data base pubblico entro 24 ore dalla scoperta.

Per quanto riguarda l'Italia, ha partecipato al Progetto Genoma Umano dal 1988 al 1994. Il coordinamento era stato affidato a Renato Dulbecco e i finanziamenti derivavano dal Cnr, per un totale di 10 miliardi di lire in sei anni.

La ragione per la fine della partecipazione italiana a questo progetto è stata dovuta alla sospensione dei finanziamenti dal 1995 in poi.

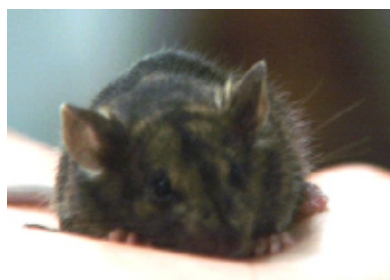
5 I progetti ancillari

Due sono i principali progetti che accompagnano il PGU:

(1) Il progetto “Human Genome Diversity Project”: progetto volto ad identificare e caratterizzare le variazioni genetiche nelle differenti popolazioni umane. L’obiettivo iniziale era di collezionare il DNA di 25 individui, che manterranno l’anonimato, per ognuno dei 500 gruppi etnici selezionati (si stima ce ne siano 7000).

I fini scientifico-culturali di questo progetto sono: - lo studio delle origini dell’uomo, dei movimenti e delle strutture sociali delle popolazioni preistoriche - lo studio della diversa suscettibilità alle malattie per discriminare quanto questo dipenda dalla genetica e quanto dall’ambiente in cui si vive. Tra le malattie di notevole interesse per lo studio si annoverano il diabete, l’ipertensione, la talassemia e l’anemia falciforme - annoverare informazioni utili nell’ambito dell’antropologia forense: le impronte di DNA sono ampiamente utilizzate in sede legale per la loro accuratezza, ma il tipo di analisi che viene fatto è suscettibile all’appartenenza etnica (quello che può caratterizzare in modo specifico un soggetto caucasico potrebbe essere un tratto molto comune in un orientale).

(2)



Il sequenziamento del genoma di altri organismi, perché lo studio di meccanismi altamente conservati tra le specie possa continuare in aree di maggior importanza per la comprensione della biologia umana.

Un dossier di Laura Barberis e Marika De Acetis, aggiornato al 10.03.2004

5.1 Variabilità genetica di popolazione

Il 99,9% del DNA è conservato tra individui diversi, quindi una differenza dello 0,1% nella composizione del DNA è sufficiente per dar luogo alla variabilità di popolazione. Alcuni studiosi spiegano l’affermarsi di un particolare “set” di geni con la teoria della deriva genetica.

Il concetto principale su cui si basa questa teoria è che nelle piccole comunità le frequenze di alcuni geni cambiano in maniera rilevante da una generazione all’altra e dopo molte generazioni è plausibile che popolazioni diverse abbiano sviluppato diverse frequenze geniche. Quindi la differenza nel genoma, la “distanza genetica”, può essere utilizzata per ricostruire le suddivisioni nella storia delle varie popolazioni.

Per tutte le coppie di popolazioni paragonate la maggior parte dei geni risulta essere in comune, e alcuni geni sono altamente conservati, mentre altri, pochissimi, variano in maniera rilevante.

Negli studi effettuati da Cavalli-Sforza e coll. risulta che la maggior divergenza genetica è presente fra i Mbuti Pigmei e gli aborigeni australiani. Se si considera la popolazione inglese, i più vicini dal punto di vista genetico sono i danesi, mentre i più lontani sono i Mbuti Pigmei e la differenza con le altre popolazioni europee è veramente minima.

Nel tentativo di disegnare l’albero genealogico della popolazione umana in base a dati genetici e molecolari, la prima divisione sembra essere stata tra gli africani e il resto del mondo, convalidando la teoria che l’uomo sia comparso per la prima volta in Africa e poi alcuni individui siano emigrati verso est, e solo più tardi si siano divisi in più popolazioni.

Secondo Cavalli-Sforza e coll. la successiva divisione sembra essere stata dei cosiddetti non-africani in australiani, quindi negli asiatici del sud-est, e solo in seguito, dalla popolazione rimanente si sarebbero divisi i caucasici, gli asiatici del nord est, quelli artici e gli indiani americani.

Per quanto riguarda l’Europa, le popolazioni che risultano avere una maggiore divergenza rispet-

to al resto del continente sono i lapponi e i sardi, entrambi caratterizzati da un lungo periodo di isolamento e da un basso numero di componenti. A seguire vengono i baschi, ritenuti i più simili alla popolazione europea originaria, e gli islandesi. Risultano essere molto simili dal punto di vista genetico gli inglesi, i tedeschi, gli svedesi, gli italiani, i polacchi e i russi. Si discostano lievemente da questo gruppo gli scozzesi e gli irlandesi.

Nel DNA si possono trovare le tracce delle pressioni climatiche e delle malattie endemiche di determinate regioni, delle espansioni preistoriche ma anche quelle più recenti, dovute allo sviluppo dell'agricoltura.

Detto ciò, paradossalmente, bisogna in realtà considerare che la diversità genetica tra gli individui di una popolazione è maggiore di quella media osservabile tra due popolazioni.

Lecture consigliate: *The History and Geography of Human Genes*, Cavalli-Sforza L.L., Menozzi P., Piazza A.; Princeton: Princeton University Press 1994.

5.2 Il sequenziamento di altri genomi

I genomi di numerosi organismi, oltre all'uomo, è stato sequenziato o è in via di sequenziamento. Nella tabella sono riassunti i passi principali compiuti.

Tabella dei genomi non umani sequenziati o in sequenziamento

6 Dove si trovano i risultati?

I risultati ottenuti dal sequenziamento del genoma umano vengono inseriti in un data base a cui tutti possono accedere.

I ricercatori che partecipano a questo progetto, nell'ambito del finanziamento pubblico, hanno l'obbligo di inserire i risultati ottenuti nel database entro 24 ore dalla scoperta.

Questi programmi sono strutturati in modo da permettere delle ricerche in base al nome del

gene, ad un tratto di sequenza, al nome della proteina e così via.

Basta immettere i dati che si conoscono, e il computer ricerca nell'archivio le informazioni note sull'argomento richiesto.

Si possono ottenere diversi tipi di informazioni: ad esempio se la stessa sequenza individuata nell'uomo è presente in altri animali, e con quale grado di omologia, se quella determinata sequenza appartiene ad un gene già identificato e da chi è stato scoperto.

Si possono ottenere le sequenze complete dei geni identificati, con la sequenza proteica corrispondente, l'identificazione delle sequenze particolari e delle sequenze consenso, cioè quelle sequenze che codificano per porzioni proteiche che rivestono dei ruoli particolari, e si possono ritrovare in diverse proteine.

Ove possibile sono fornite mappe e implicazioni biologiche della sequenza e la bibliografia sottoforma di link ipertestuale.

6.1 La gene bank

Ecco un esempio delle informazioni che si possono ottenere dalla gene bank.

Volendo ottenere delle informazioni relative all'emoglobina, la proteina che trasporta l'ossigeno nel nostro organismo, è sufficiente digitare la parola "hemoglobin" nella casella di ricerca della GenBank.

Si possono fare molti tipi di ricerca. Supponiamo di voler fare la ricerca per "LocusLink" che fornisce informazioni sulla posizione nel genoma del gene che codifica per questa proteina.

Quello che compare è una tabella in cui sono elencati tutti i geni codificanti per le varie forme dell'emoglobina, poiché ne esistono diverse.

Per ogni gene, o pseudogene (cioè un fossile genetico che presenta la sequenza codificante dell'emoglobina, in questo caso, ma che non è in grado di dirigerne la sintesi poiché manca di regioni importanti per questa funzione) sono indicati il numero con cui sono identificati in banca dati (LocusID - ID è acronimo di identification,

identificazione), l'organismo da cui è stata isolata la sequenza (Org, ad esempio Hs sta per Homo sapiens), l'acronimo con cui viene indicato (Symbol), la descrizione per esteso (Description), la posizione sul cromosoma (Position) e dei link per altre informazioni (links).

6.2 I siti web

I principali siti che contengono informazioni sul PGU e sui frammenti di DNA sequenziati sono elencati nella tabella.

Link ai siti sul Progetto Genoma Umano

7 Questioni bioetiche e aspetti commerciali

Per stabilire i principi etici alla base delle ricerche del Progetto Genoma Umano è stata fondata l'ELSI (acronimo per ethical, legal, social issue).

Gli scopi di questa commissione internazionale sono l'educazione e la compilazione di linee guida per la conduzione della ricerca in campo genetico, la salvaguardia della privacy degli individui che forniscono il DNA per lo studio della variabilità genetica, le questioni concernenti il razzismo, l'identificazione dei possibili utilizzi, positivi e negativi, dei dati genetici, l'identificazione degli utilizzi commerciali e l'impatto delle nuove conoscenze nella medicina forense.

7.1 A chi appartiene il genoma umano?

L'interesse di molte società private per il Progetto Genoma Umano, e per i progetti ancillari, deriva dalla possibilità di sfruttare economicamente il sequenziamento del DNA, previo ottenimento del brevetto dei frammenti.

Dal 1981 al 1995 il numero di brevetti su sequenze di DNA concessi nel mondo è stato di 1175 di cui solo il 17% proprietà di enti pubblici (Nature 380 387-388 1996), e il numero di brevetti richiesti al 1996, includendo cioè anche quelli

non ancora concessi, era di 6105 (Nature 384 608 1996).

Fino all'inizio del 2001 sono stati richiesti oltre 3 milioni di brevetti; la sola Celera Genomics ne ha richiesti nel 1999 ben 6500, e gli introiti previsti devono essere ben elevati se si considera che il budget stimato per il Progetto Genoma Umano da Greg Venter all'interno della sua compagnia era di 300 milioni di dollari.

Il Comitato Nazionale di Bioetica ha espresso un parere sulla brevettabilità del genoma umano sottolineando l'importanza di una "... equa condivisione fra tutti gli Stati dell'accesso all'applicazione, utilizzazione e sfruttamento dei nuovi risultati conseguiti dalla ricerca sul genoma umano...", tuttavia non chiarisce "...se le nuove conoscenze acquisite dalla ricerca siano a disposizione di tutti - quale bene comune dell'umanità - al quale ognuno può accedere ...".

I brevetti vengono concessi in base a quattro caratteristiche. L'inventore deve aver identificato una sequenza genetica nuova, deve specificare il ruolo di questa sequenza e la funzione di questa in natura, deve descriverla in modo che un altro specialista del campo possa utilizzarla per i propri scopi.

I prodotti della natura non sono patentabili, e per quanto riguarda il DNA, il brevetto può essere concesso solo su sequenze isolate, purificate o modificate in maniera tale da non esistere in natura.

Un problema rilevante è che per molte sequenze è in corso l'attribuzione del brevetto, ma queste sono già presenti in banca dati, pertanto il loro uso sottoporrà in futuro gli utilizzatori al pagamento dei diritti ai detentori del brevetto.

In Europa la direttiva 98/44/CE dell'UE stabilisce che non si può brevettare la sequenza di un gene in quanto tale, ma solo una sua specifica applicazione.

In realtà anche questo concetto viene applicato in maniera discutibile, poiché per ottenere il brevetto di sequenze ignote le compagnie definiscono l'applicazione sulla base di simulazioni al computer, che sono ovviamente ancora molto lontane dal poter descrivere il complesso

funzionamento biologico del nostro genoma.

Per quanto riguarda il parere dell'HUGO, si oppone alla concessione di brevetti per brevi sequenze di cui la funzione non è chiara e ribadisce il concetto che solo l'immediata pubblicazione delle sequenze scoperte permette la libera diffusione delle conoscenze, alla base della cooperazione internazionale fondamentale per l'avanzamento scientifico.

I brevetti consentono di sfruttare le invenzioni senza nascondere i dati e l'utilizzo esclusivo è relativo ai 20 anni successivi alla concessione del brevetto, e il motivo per cui vengono concessi è l'incentivazione ad investire nella ricerca, con la garanzia di avere un tornaconto.

Questa possibilità di sfruttamento economico ha sicuramente dato un importante impulso per la crescita della ricerca scientifica, grazie al coinvolgimento di imponenti capitali da parte di compagnie private.

I rischi connessi al rilascio dei brevetti in campo genetico sono però molteplici.

Innanzitutto possono interferire con lo sviluppo di mezzi terapeutici e diagnostici a causa dei costi associati all'uso dei dati coperti da brevetto.

Inoltre i prodotti derivanti da una sequenza brevettata hanno costi maggiori perché il produttore deve ammortizzare gli elevati costi della richiesta di brevetto.

Infine mettono in luce un problema chiave in questo settore: se è possibile brevettare un elemento appartenente alla natura, questo può significare che un individuo può diventare possessore di parte di un altro individuo?

7.2 Un ritorno dell'eugenetica?

La disponibilità della completa sequenza del genoma e la conseguente possibilità di identificare le persone portatrici di alterazioni genetiche potrebbe avere delle serie conseguenze.

Le compagnie assicuratrici nonché i datori di lavoro potrebbero utilizzare in modo non etico le informazioni desumibili da un futuro

“passaporto genetico”.

Il comportamento etico da osservare è che l'analisi genetica venga fatta solo su esplicita richiesta dell'interessato e i risultati devono essere coperti da privacy.

L'8 febbraio 2000 il Presidente Clinton ha firmato un ordine esecutivo che proibisce l'utilizzo dell'informazione genetica nelle procedure che portano all'assunzione al lavoro, alla promozione lavorativa, alla stipulazione di contratti assicurativi.

Il documento stabilisce che non possono essere fatte discriminazioni in campo lavorativo e assicurativo in base alla probabilità di sviluppare delle malattie genetiche.

L'unica eccezione riguarda la predisposizione genetica a sviluppare delle malattie se esposti a particolari ambienti di lavoro, ma questo nel rispetto del consenso informato che il lavoratore deve avere se si appresta a lavorare in condizioni per lui nocive.

7.3 Il parere del CNB italiano

Il CNB (Comitato Nazionale per la Bioetica) italiano ha espresso il seguente parere sul Progetto Genoma Umano.

Il CNB sottolinea che il Progetto Genoma Umano è un'impresa scientifica di grandi dimensioni e rilevanza, che pur non comportando delle nuove problematiche ha riportato alla luce dei problemi etici pre-esistenti al PGU, amplificandoli e concentrandoli.

Il CNB raccomanda di adoperarsi affinché le nuove scoperte scientifiche in questo ambito, che hanno un'importanza fondamentale nell'avanzamento della conoscenza, non portino a descrivere l'uomo solo in base ad una visione meramente scientifica, e soprattutto non siano un mezzo di discriminazione.

Propone come metodica di condotta la continua riflessione etica, anche durante la progettazione della ricerca, rimettendosi in questo caso all'etica professionale del ricercatore, da attuarsi anche facendo delle indagini sull'impatto sociale e giuridico delle nuove scoperte.

Inoltre raccomanda come utile strumento per ovviare alla distorsioni delle notizie scientifiche e all'uso improprio delle conoscenze scientifiche, l'informazione del pubblico.

Infine il CNB ritiene di fondamentale importanza la gestione riservata dei dati personali derivanti dall'analisi genetica, in particolare per l'uso improprio che ne potrebbe derivare sia per l'assunzione ai fini lavorativi, sia nella stipulazione di contratti assicurativi, sia nei tentativi di modificazione dell'assetto genetico di alcune popolazioni (eugenetica positiva e negativa).