

BIOINFORMATICA

Între biologie și tehnologie

ENRIQUE M. MURO

Traducere de Ionela Polixenia Matei

CUPRINS

Capitolul 0

Prima citire a genomului uman	7
Am ajuns pe Lună!	7
Dar de ce?	9
Proiectul Genomul Uman	12
Spune-mi, bioinformatică, cine ești?	17

Capitolul 1

Cartea vieții	19
În ce a constat citirea ei?	19
Literatura critică	24
<i>De ce heterocromatina a îngreunat prima citire?</i>	31

Capitolul 2

Pilonii bioinformaticii	33
Scurt istoric al cercetării ADN-ului	34
<i>Numărul de cicluri de diviziune celulară necesar pentru a obține un adult pornind de la un zigot</i>	38
<i>Comparație între dimensiunile genomului uman și al lui „Don Quijote de la Mancha” de Cervantes</i>	43
Proteinele	43
<i>De ce fiecare codon ce codifică un aminoacid este format din trei nucleotide?</i>	45
Evoluția	49

Capitolul 3

Bioinformatică, de unde vii?	53
Geneza computerelor	54
Computerele se apropie de biologie	56
<i>Un microprocesor pentru fiecare neuron</i>	56
Bioinformatica se afirmă ca disciplină	60

Capitolul 4

Secvențe și alinieri	69
Secvența reprezentată în format FASTA	70
<i>Exemplu de format FASTA</i>	72
Alinierea secvențelor	79
<i>Analogie între limbi și secvențe de proteină</i>	82
<i>Un exemplu al distanței Hamming</i>	85
Matrice de substituție	86
Indel	88
<i>Cine a fost primul, oul sau găina?</i>	89
<i>Cât valorează un indel?</i>	90
Algoritmi de aliniere	92
<i>Ce tip de algoritm trebuie să aleg?</i>	93
<i>Este adevărat că programarea dinamică reduce timpul de calcul?</i>	98

Capitolul 5

Relații între specii: arborii filogenetici	103
Clasificarea vieții pe Pământ	103
Relațiile evolutive între diferitele specii	104
<i>Un exemplu simplificat al construcției unui arbore filogenetic</i>	106
Dezvăluind misterul vieții pe Pământ	111

Capitolul 6

Bioinformatica ne ajută să ne cunoaștem mai bine	119
Cine suntem?	119
De unde venim?	120
<i>Câte generații ne separă de Neanderthal?</i>	128
Încotro mergem?	134
Lecturi recomandate	143

Capitolul 0

Prima citire a genomului uman

Am ajuns pe Lună!

A fost o noapte alb-negru, atât de lungă încât tiparele activității neuronale se referă încă la amintirile celor care au trăit-o. A fost noaptea cifrelor care nu s-au oprit din creștere pe măsură ce se întuneca. Așa s-a întâmplat că, în țara mea, aproximativ 20 de milioane de oameni dintr-o populație care nu ajungea la 34 de milioane și-au petrecut noaptea împreună în familie, la casele prietenilor sau oriunde ar fi fost unul dintre cele abia 3 milioane de televizoare disponibile. Se estimează o audiență de 530 de milioane de spectatori în întreaga lume, pentru o producție care implica peste o mie de oameni și care a costat aproximativ 11 milioane de dolari la acea vreme. În mod curios, una dintre cele trei antene cu diametrul de 26 de metri pe care NASA le-a instalat pentru a recepționa semnalul și a-l putea retransmite pe întreaga planetă era situată în provincia Madrid, mai exact în Fresnedillas de la Oliva.

Grijile legate de somn și de faptul că trebuia să se lucreze a doua zi erau acum mai puțin importante: ceea ce urma să se întâmple avea să recompenseze noaptea albă. În 21 iulie 1969, comandantul misiunii spațiale Apollo 11 avea să pună piciorul pe Lună pentru

prima dată, lăsând în cărțile de istorie acea amprentă celebră la sud de Marea Liniștii. Bărbatul care purta acea cizmă se numea Neil Armstrong, iar ceasul din camera mea arăta aproape patru dimineața. Câteva momente mai târziu a rostit sintagma „Un pas mic pentru om, un salt uriaș pentru omenire”.

Toate acestea au avut loc în timpul Războiului Rece dintre Uniunea Sovietică, acum dispărută, și Statele Unite ale Americii, confruntare care, în cele din urmă, nu a izbucnit într-un război mondial, dar ale cărei tensiuni socio-politice continue au ținut întreaga planetă în suspans. Cele două superputeri s-au întrecut, printre altele, pentru a fi primele care aterizează pe Lună, la bordul unei navete spațiale cu echipaj. A fost un conflict lung care a lăsat în urmă o serie de fapte și realizări mari din partea ambelor părți. Sovieticii au avut un mare avantaj inițial, deoarece au fost primii care au ajuns pe Lună, la bordul unei navete spațiale fără echipaj, în misiunea spațială Luna 2, în data de 13 septembrie 1959. În plus, astronautul Yuri Gagarin, de asemenea sovietic, a fost primul care a reușit să orbiteze în jurul Pământului în data de 12 aprilie 1961. Doar trei săptămâni mai târziu, Alan B. Shepard a fost primul american care a mers în spațiu, chiar dacă a reușit să urmeze doar o traiectorie suborbitală. Câți copii născuți în acei ani s-or fi numit Yuri! Ca răspuns la insultă, John F. Kennedy a anunțat într-un discurs celebru (pronunțat la Universitatea Rice, Houston, în 1962) intenția țării de a avea unul dintre concetățenii săi care să ajungă pe Lună și apoi să-l aducă înapoi, toate acestea înainte de sfârșitul deceniului. Nu a fost Kennedy cel care a văzut cursa spațială încheiată cu succes, ci Nixon, în condițiile în care primul a fost asasinat la puțin peste un an de la declarație.

Naveta spațială care a trecut de provocarea lui Kennedy a fost pilotată de Neil Armstrong, Edwin Buzz Aldrin și Michael Collins. Collins nu a putut să pășească pe pământul lunar, deoarece a rămas la comanda modulului principal al navetei. Datorită lui și unei mari echipe de oameni de știință și ingineri, restul echipajului a reușit să exploreze satelitul, la bordul modulului lunar *Eagle*. Ei au adus pe

Figura 1. Prima pagină
a *Washington Post*
din 21 iulie 1969.

Pentru prima dată,
doi oameni puseseră
piciorul pe Lună.



Pământ 21,5 kg de mostre pentru analize ulterioare, au efectuat o serie de experimente științifice și au lăsat acolo câteva obiecte, cum ar fi oglinda reflectorizantă, despre care ministrul spaniol al științei și astronautul Pedro Dunque au vorbit în iulie 2018, ca dovadă pentru a combate argumentele contestatarilor. Mai exact, oglinda reflectorizantă este folosită pentru a măsura cu precizie distanța de la Pământ la Lună.

Apollo 11 a fost prima misiune pilotată care a sosit pe Lună la bordul rachetei Saturn V, urmată de încă cinci misiuni. Din acel moment și până pe 14 decembrie 1972, un total de doisprezece bărbați au aterizat pe Lună. Ulterior, misiunile de acest tip au fost suspendate, deși explorarea spațiului nu a fost. Astăzi, toți cetățenii beneficiază de progresele rezultate din această dezvoltare aerospațială. Nu ne-a permis doar să ne îmbogățim cunoștințele în științele de bază, cum ar fi fizica, matematica, biologia, medicina și așa mai departe; dezvoltarea a implicat și științele aplicate, cum ar fi informatica, electronica, robotica, teledetecția, meteorologia, navigația, sateliții, comunicațiile...

Dar de ce?

Aterizarea omului pe Lună a fost un spectacol tehnologic extraordinar fără precedent și, în același timp, a fost destul de simplu de înțeles. Importanța a ceea ce se întâmpla a fost evidentă, în ciuda

complexității științifice și tehnologice subiacente. Pentru prima dată, ființa umană nu a privit singură și absorbită la fereastra universului. De data aceasta, se afla de cealaltă parte a imensității necunoscute a unei nopți înstelate, pe suprafața Lunii, sprijinindu-se pe prima treaptă a unei scări despre care, în viitor, dorim să ne conducă în cele mai îndepărtate locuri ale cosmosului. Acel prim pas al lui Neil Armstrong pe Lună și amprenta piciorului său simbolizează împlinirea unui vis, despre care nu știm de când datează, în cei puțin sub 4000 de milioane de ani de evoluție a vieții pe Pământ sau în ultimii 6,7 milioane de ani, din momentul în care specia noastră s-a diferențiat de cimpanzei. Este destul de probabil ca în acest interval de timp și alte specii deja dispărute să fi avut același vis, inclusiv hominizii arhaici precum Omul de Denisova sau de Neanderthal. Nu există nicio îndoială că a meritat să stai treaz în acea noapte.

Realizarea acțiunii desfășurate de echipajul misiunii Apollo 11 depășește efortul de a îmbunătăți societatea prin intermediul progresului științifico-tehnologic sau dorința de a merge mai departe, de a explora. Ființele umane au fost întotdeauna preocupate să reflecteze asupra unei serii de întrebări tipice pe care diferite generații și le-au pus, una după alta, de-a lungul istoriei umanității. Sunt întrebări care apar în mod recurent în artă și în știință, ca în imaginea prezentată în tabloul lui Paul Gauguin: De unde venim? Cine suntem noi? Încotro mergem? Nu am găsit și nu știm dacă într-o zi vom găsi răspunsuri satisfăcătoare și complete. Am încercat cu tot ceea ce avem la dispoziție, dar pentru moment suntem departe de a rezolva aceste îndoieli. Cu toate acestea, știm ceva. Nu știm cine suntem, dar știm că suntem formați dintr-o combinație de atomi de diferite elemente ale tabelului periodic în proporții diferite: carbon (C), hidrogen (H), oxigen (O), azot (N), fosfor (P) și așa mai departe. Datorită fizicii nucleare, suntem conștienți de faptul că atomii mai grei se formează din atomi mai ușori în interiorul stelelor, la presiune și temperatură foarte ridicate. Dintre primii, atomii cu o greutate și mai

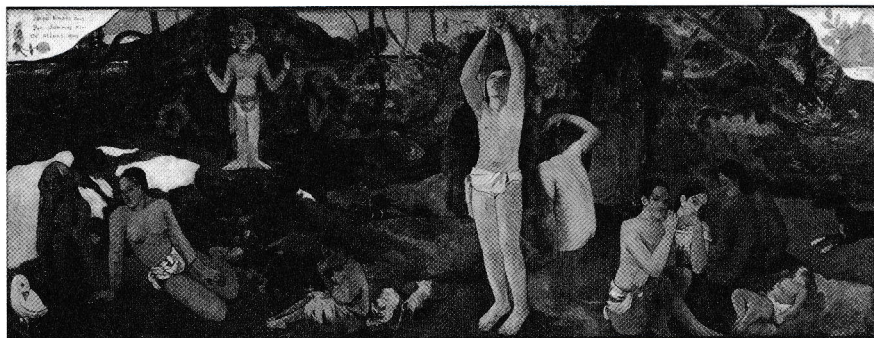


Figura 2. Pictură postimpresionistă de Paul Gauguin, pictată în Tahiti în 1897. În colțul din stânga sus se poate citi titlul: *D'où venons nous / Que sommes nous / Où allons nous* (în franceză, „De unde venim? Cine suntem? Încotro mergem?“). Tabloul este inspirat de întrebările pe care un tahitian le-ar pune unui străin, dându-le o conotație filosofică.

mare, începând de la fier (Fe), se formează în condiții și mai speciale, și anume în supernove, care sunt explozii stelare. Cu alte cuvinte, noi, ființele umane, suntem alcătuite practic din atomi care într-un timp îndepărtat și-au avut originea în stele îndepărtate, în locurile îndepărtate ale universului. Minunat! Fără îndoială celebra maximă a cărții Genezei „Pământ ești și în pământ te vei întoarce”, conform acestei perspective, este adevărat; în concluzie, suntem praf de stele, iar atomii din care suntem făcuți sunt veșnici.

Deși, în realitate, o ființă vie este mult mai mult decât un sistem complex și nu este reductibilă la suma atomilor care o compun, atomi care se combină prin interacțiuni chimice, dând viață moleculelor lungi și biomoleculelor, precum proteinele. Acestea din urmă formează structuri tridimensionale care, asemenea pieselor de Lego, se cuplează pentru a forma construcții din ce în ce mai complexe care, la rândul lor, se integrează cu altele și constituie organismul. Nu știm cine suntem, dar știm din ce suntem făcuți și de unde provine materia care ne formează.

Observarea și explorarea universului sunt surse de cunoaștere, iar aterizarea omului pe Lună a fost un eveniment istoric, un stimul inegalabil care a servit drept stimulent pentru câteva generații pentru a continua să caute răspunsuri, încercând să înțeleagă ceea ce înconjoară planeta noastră și întorcându-și privirea spre exterior. În același timp, această piatră de hotar a impulsionat dezvoltarea tehnologică și a încurajat conștientizarea beneficiilor pe care știința, în general, și știința aerospațială, în special, le aduc cetățenilor.

Proiectul Genomul Uman

Abia treizeci de ani mai târziu s-a atins un alt reper științific care a constituit și unul dintre cele mai semnificative progrese din istoria omenirii, datorită căruia intrăm direct în subiectul tratat în această carte. Nu a fost un eveniment la fel de mediatizat ca aterizarea omului pe Lună, deoarece scopul nu era defel evident. S-a investigat acolo unde cetățenii nu pot vedea direct cu ochii, într-un loc care nu este intuitiv și este accesibil mai întâi prin gândire și apoi prin visare; un loc unde ne-a călăuzit gândul lui Socrate, reflectat în acel aforism înscris pe frontispiciul templului lui Apollo din Delphi: „Cunoaște-te pe tine însuți”. S-a cercetat ceea ce este mic, ceea ce este în interiorul nostru, unde nici măcar cele mai sofisticate microscopie nu pot ajunge. Mă refer, desigur, la prima lectură a genomului uman, a materialului genetic care a călătorit de mii de ani prin viață, transportat de la specie la specie, ajungând până la noi.

Informația oficială a realizării primei schițe a genomului uman datează din 26 iunie 2000. O noutate de această amploare a trebuit să fie dezvăluită lumii întregi într-un mod semnificativ, astfel încât, cu această ocazie, divulgatorii au fost Bill Clinton și Tony Blair, președintele Statelor Unite ale Americii, respectiv premierul Regatului Unit. Imediat, televiziunile și presa mondială au făcut ecou știrii.

Figura 3. Prima pagină
a publicației *El País*
din 27 iunie 2000.
Cu o zi înainte,
Bill Clinton și Tony Blair
anunțaseră prima schiță
a genomului uman.



Adevărații protagoniști ai acelei acțiuni au fost însă oamenii de știință care au realizat-o, reprezentați de liderii echipei de cercetători, Francis Collins și Craig Venter.

Au reușit să ne citească genomul pentru prima dată! Opinia publică nu putea înțelege, în complexitatea ei, valoarea a ceea ce se anunța, dar era clar că tocmai se întâmplase ceva important, în condițiile în care, în acei ani, știrile referitoare la știință, în general, nu atrăgeau atenția mass-media și cu atât mai puțin pe cea a televiziunii. S-a emis ipoteza că acest nou progres științific ar putea deveni calul de bătaie al medicinei în lupta cu numeroase boli, în special împotriva cancerului. Și așa a fost; în țările dezvoltate s-a înregistrat o creștere a ratei de supraviețuire, în cazul bolilor grave, precum și a speranței de viață a populației. Citirea genomului nostru a fost esențială pentru a obține astfel de îmbunătățiri în sănătatea cetățenilor. Pe de altă parte, ne-am sporit cunoștințele despre științe de bază, cum ar fi biologia, medicina și așa mai departe. Cu alte cuvinte, înțelegem mult mai

bine cum funcționează organismul nostru. Așa cum s-a întâmplat în cazul aterizării omului pe Lună, citirea genomului uman a încurajat dramatic dezvoltarea tehnologică pe care am experimentat-o în ultimele decenii, ale cărei consecințe au depășit limitele biomedicinii. Dacă descoperirile oferite de fizică la începutul secolului XX au avut o importanță enormă în dezvoltarea științifică și tehnologică a societății noastre, trebuie să recunoaștem același rol principal în citirea genomului nostru, care a avut loc la începutul secolului XXI. Astăzi există titluri frecvente în presă care au ca protagonist genomul și acest lucru demonstrează o schimbare recentă în societatea noastră.

Proiectul Genomul Uman a început oficial în 1990, dar dezbaterea asupra implementării lui datează deja de la începutul deceniului precedent. Guvernul SUA a alocat 3000 de milioane de dolari tuturor NIH-urilor (Institute Naționale de Sănătate), bani care au venit din sectorul public. Costul final a fost ceva mai mic, egal cu aproximativ 2700 de milioane de dolari. În plus, proiectul a fost estimat să dureze cincisprezece ani, dar s-a încheiat cu doi ani mai devreme. Conducerea proiectului i-a fost încredințată lui **James Watson** (n. 1928), laureat al Premiului Nobel și unul dintre descoperitorii structurii de dublu helix a ADN-ului. Demisia sa din această funcție în 1992 a fost categorică și a fost produsă de divergențe legate de posibilitatea ca genele să poată fi brevetate. Watson și-a exprimat clar poziția: „Națiunile lumii trebuie să înțeleagă că genomul uman aparține umanității și nu națiunilor”. El a apărat cunoștințele și progresul științific care ne aparțin tuturor. Încă nu exista o prevedere legală în acest sens, iar repercusiunile socio-economice puteau fi enorme. Să ne imaginăm, de exemplu, că secvența genei care produce insulină (o proteină) ar fi brevetată. Apoi, orice produs derivat din secvența acestei gene, cum ar fi cel pe care diabeticii trebuie să-l ia prin injecție, ar fi supus brevetului. Ar trebui să existe mii de brevete, cel puțin unul pentru fiecare genă. Un alt exemplu izbitor se referă la gena *CFTR* care provoacă fibroza chistică.

Dacă gena ar fi brevetată legal și o echipă de oameni de știință ar încerca să găsească un remediu pentru această boală, echipa științifică ar trebui să suporte costul respectiv pentru a putea face cercetări asupra acestuia. Ar fi o catastrofă pentru că ar crește costurile cercetării și asta ar presupune o încetinire a progresului medical. După Watson, conducerea *Proiectului Genomul Uman* a trecut la **Francis Collins** (n. 1950) până la încheierea acestuia. Collins a fost precedat de reputația sa de *vânător de gene*, întrucât, în trecut, colaborase la identificarea genelor care provoacă boli grave precum fibroza chistică, boala Huntington, neurofibromatoza și așa mai departe.

Trebuie avut în vedere că, atunci când proiectul era la început, obținerea secvenței unei singure gene putea dura mulți ani de muncă pentru o echipă de cercetare. În 1988, erau cunoscute doar aproximativ 1400 de gene și foarte puține soiuri genetice responsabile de boli. În prezent, însă, au fost identificate mii de variații genetice asociate patologiilor umane; în plus, este cunoscută secvența completă a genomului nostru și au fost localizate puțin sub 20 000 de gene care codifică proteinele. Este foarte important de subliniat că toate informațiile care decurg din *Proiectul Genomul Uman* și adnotările ulterioare, obținute de la mii de echipe de cercetare de pe tot globul, sunt disponibile publicului, în baze de date care pot fi accesate de oricine printr-o simplă conexiune la internet.

Deși proiectul a început la INS din SUA, multe institute și centre de cercetare și-au unit eforturile, creând o sinergie de cooperare internațională de care să beneficieze omenirea pentru totdeauna. Dintre toate instituțiile din afara Statelor Unite, trebuie evidențiat Institutul *Wellcome Trust Sanger Institute* din Regatul Unit.

În 1998, a existat o amenințare pentru proiectul public, o inițiativă din sectorul privat sub conducerea lui **Craig Venter** (n. 1946). Acesta din urmă a comunicat că va putea finaliza proiectul în trei ani și, în acest scop, a înființat o firmă, Celera, al cărei nume era deja un afront

la adresa proiectului public, întrucât se referea la rapiditatea cu care intenționau să lucreze. Dacă ar fi respectat termenele, previziunile consorțiului public ar fi fost anticipate cu patru ani. Pe lângă faptul că proiectul original ar putea pierde concurența pentru atingerea obiectivelor propuse, erau în joc brevetele asupra genelor pe care Celera pretindea că le obține. Acționarii Celera investiseră 300 de milioane de dolari și, în schimb, sperau la o rentabilitate economică suculentă. Cursa pentru secvențierea genomului uman începuse, la fel ca și cursa spațială cu decenii înainte. În cele din urmă, apariția bruscă a Celera a devenit un mare stimulent pentru a finaliza proiectul mai devreme decât era prevăzut.

Strategia Celera s-a bazat pe tehnologia de secvențiere (în engleză *shotgun sequencing*) care fusese anterior utilizată în mod satisfăcător cu genomuri bacteriene, cu trei ordine mai mici ca mărime. La rândul lor, membrii consorțiului public, pentru a împiedica Celera să breveteze părți ale genomului uman pe care ei înșiși le-au secvențiat, au pus zilnic la dispoziția lumii întregi toate secvențele obținute.

În martie 2000, cu câteva luni înainte de știrea primei schițe a genomului uman, s-a luat o decizie legislativă privind brevetele: secvențele obținute din genomul uman nu pot fi brevetate. Acest lucru s-a întâmplat cu câteva luni înainte de știrea primei schițe a genomului uman, pe 26 iunie 2000. A fost nevoie de doar câteva luni, până în februarie 2001, pentru a fi publicate articolele științifice care descriu rezultatele. În cele din urmă, la 14 aprilie 2003, proiectul a fost considerat finalizat. Nu a existat un singur câștigător final, având în vedere că atât consorțiul public, cât și cel privat au fost creditate cu îndeplinirea obiectivelor, în timp ce citirea genomului uman a rămas disponibilă pentru totdeauna umanității. Dar toate acestea ar fi fost imposibil de realizat fără ajutorul protagonistului incontestabil al acestei cărți: bioinformatica.