

cat | esp | en

talent

dimecres, 24 febrer 2010

TOP

actualitat

fronteres

ciència per a presidents

MÓN

corresponsals

bitàcoles

OFF CIÈNCIA

tribunes

la tira

DIVULGADOR

entrevistes

perfils

conferències

GALERIA

fotogaleries

documentals

infografies

PÒSTIT

recomanacions

crítiques

agregador cultural

SISTEMA

entorns

recursos

convocatòries

INICI > TOP > Actualitat > Dades biològiques, de Babel a BIANA

Actualitat

BIOINFORMÀTICA



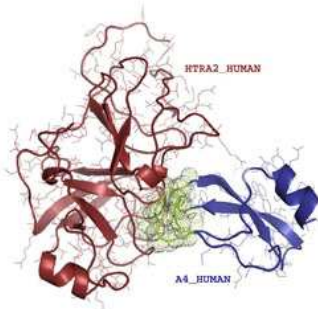
Dades biològiques, de Babel a BIANA

Una nova plataforma integra diferents bases de dades fins ara disperses

El desgavell de dades biològiques, fins ara disperses a diferents bases de dades, arriba a la seva fi. Un nou programa d'accés lliure, BIANA, s'ha posat a la pell dels investigadors i les ha aplegat per facilitar-los la recerca. Amb aquesta nova eina, de segell català, els científics tindran al seu abast una gran quantitat d'informació, podran trobar noves relacions entre diferents elements biològics, fer prediccions i identificar noves dianes terapèutiques.

PATRICIA MORÉN | 22 DE FEBRER DE 2010

La recerca amb dades biològiques serà més fàcil i àgil a partir d'ara. Una nova plataforma, BIANA (acrònim de *Biologic Interactions and Network Analysis*), ho ha fet possible, després d'aplegar una quantitat ingent de dades disperses en diverses fonts i emmagatzemades amb diferents formats. Els investigadors que consultin la nova plataforma tindran al seu abast totes les bases de dades integrades. Podran manipular-les en funció del seu interès de recerca, inferir noves interaccions entre proteïnes i trobar noves dianes terapèutiques, entre altres usos. La revista *BMC Bioinformatics* ha descrit el nou programari.



La nova eina ha estat íntegrament desenvolupada pel Laboratori de Bioinformàtica Estructural, dirigit per Baldomero Oliva, dins del Grup de Recerca en Informàtica Biomèdica (GRIB), de la Universitat Pompeu Fabra i l'Institut Municipal d'Investigació Mèdica (UPF-IMIM). El projecte neix d'una plataforma anterior del mateix grup, PIANA (*Proteins Interactions and Analysis*), però amb més aplicacions. És accessible a tots els investigadors, de fàcil ús i disponible sota llicència GNU GPL (*General Public License*) al web de la plataforma. Un valor afegit de BIANA és que utilitza la interfície gràfica *Cytoscape*, per gestionar les dades de manera fàcil i interactiva.

"Es tracta d'un sistema obert a qualsevol usuari, que pot inserir les dades biològiques que desitgi al sistema i integrar-les a les d'altres bases de dades, en lloc d'utilitzar tantes bases concretes", explica Javier García, un dels investigadors del GRIB. Amb Emre Güney, també del grup, ha estat implicat directament en el desenvolupament del programari. Però, on rau la importància d'aquest desenvolupament bioinformàtic? Mirem enere.

Context de partida

Fins ara, l'emmagatzematge de dades biològiques s'ha dut a terme en tres nuclis principals del món: al Japó; als Estats Units, a través del *National Institutes of Health* (NIH); i a Europa, a través de l'*European Molecular Biology Laboratory* (EMBL), de Heidelberg (Alemanya), i el seu node, l'*European Bioinformatic Institute* (EBI), situat a Hinxton (Gran Bretanya). Tots ells disposen de bases de dades diferenciades.

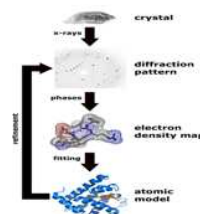
Als anys 80-90 van començar a proliferar els treballs científics de seqüenciació de gens i proteïnes. Els resultats es van anar emmagatzemant de diferents maneres i a diferents llocs. A Europa, *Swiss-Prot* (Suïssa) va ser la primera base de dades ben organitzada, fins que li va prendre el relleu *Universal Protein Resource* (UniProt), que va fusionar les dades de la base de dades suïssa i d'altres. I a tot el món en van sorgir més, com *BioGRID*, del Canadà; *Database of Interacting Proteins* (DIP) i *GenBank*, dels Estats Units; *Human Protein Reference Database* (HPRD), de l'Índia; o el *Munich Information Center for Proteins Sequencing* (MIPS), d'Alemanya. I la llista no acaba aquí.

Les diferents bases de dades contenen gran quantitat

Totes reuneixen una gran quantitat d'informació d'elements biològics. L'estudi de l'estructura de les proteïnes és una feina feixuga que pot trigar entre un i dos anys, mentre que la seva

Relacionat

Mig segle de proteïnes a través del vidre



Es compleixen 50 anys de l'anàlisi de l'estructura molecular de les proteïnes en tres dimensions

Butlletí electrònic



Vols rebre el nostre butlletí?

> apunta-t'hi

Links relacionats

Relacionats secció

Identifiquen un gen que causa problemes de moviment en ratolins

S'ha identificat una mutació que intervé en l'expressió clínica de la migranya

Una mutació genètica detectada en deu dies permet diagnosticar una malaltia

Una mutació genètica és la responsable de l'envelliment prematur

Una nova base de dades permetrà comparar i buscar gens amb major detall

La proteïna REST juga un paper clau en el procés evolutiu

Descobrixen els indicis genètics de l'autisme

d'informació, però en diferents codis i formats

seqüenciació és cada cop més ràpida. Per aquest motiu, aquestes bases de dades contenen informació de l'estructura de prop de 50.000

proteïnes, de més de 500.000 seqüències amb la seva funció (a UniProt) i de més de 10 milions de seqüències de DNA d'organismes vegetals, animals i humans, informa Oliva. El problema és que aquesta gran quantitat d'informació s'ha guardat en diferents codis i formats, a més de en diverses bases de dades. Tot plegat sembla la Babel de les dades biològiques.

Aportacions de BIANA

BIANA unifica protocols i criteris d'anàlisi d'informació de les dades biològiques fins ara disperses. Filant més prim, un dels seus trets diferencials respecte de les bases de dades esmentades és que permet crear xarxes d'interacció per a totes les proteïnes emmagatzemades, una tasca que ja duia a terme PIANA. A més, explica Oliva, obre la porta a l'exploració de tot tipus de reaccions, com ara la relació entre un enzim i un altre que pertanyen a la mateixa ruta metabòlica.

L'eina també permet fer prediccions basades en dominis coneguts de proteïnes. Això és possible perquè la base de dades UniProt, així com d'altres, no s'ha limitat a emmagatzemar seqüències senceres de proteïnes; també han acumulat dominis, és a dir, parts de proteïnes que formen un plegat propi, amb una funció pròpia i presents a moltes espècies, perquè es mantenen al llarg de l'evolució. Per tant, quan no es disposa d'una proteïna sencera, aquesta es pot inferir a partir de les funcions dels dominis. "Amb una simple integració de la informació d'aquestes bases de dades podem fer prediccions basades en dominis", emfatitza Oliva.

BIANA també pot identificar i corregir errors de la informació emmagatzemada a les anteriors bases de dades. Inclús permet que l'usuari la rastregi fins arribar a l'experiment real amb el qual es va trobar un determinat gen.

Cap al rànquing de dianes

La nova eina s'està utilitzant per estudiar malalties com l'Alzheimer, la diabetis, els aneurismes i el càncer. En el cas dels tumors, l'eina pot ser especialment útil, donat que les bases de dades contenen molta informació que es pot creuar. Gràcies a aquesta plataforma ja s'han trobat proteïnes relacionades amb la malaltia d'Alzheimer, que interaccionen amb proteïnes de la malaltia de la diabetis i que poden ser dianes potencials de tractaments farmacològics.

De fet, una de les principals utilitats d'aquest programa és que en un futur permetrà trobar noves dianes terapèutiques. Per aquest motiu, els investigadors treballen avui en la construcció d'un altre programa, NETSCORE, que permeti fer rànquings de dianes. Així, un cop s'identifiquin els candidats (imaginem-ne 1.000), el sistema els assignaria una probabilitat i els ordenaria de més a menys probable. Així, els investigadors aniran més orientats en la seva recerca i, en lloc d'estudiar un miler de candidats, només n'hauran d'analitzar uns quants.

Pel que fa a l'impacte de BIANA a la recerca, Oliva creu que, amb aquesta eina, "Fleming no s'ha acabat". Per descomptat, les troballes *by serendipity* se seguiran produint. Tanmateix, el programari permetrà ara que els resultats 'es disparin' quan això passi, donat que permet tractar la informació que es troba i assenyalar un camí adequat l'investigador.

Tòpics de l'article

[BIANA](#) [UPF](#) [Cytoscape](#) [GRIB](#) [IMIM](#) [Baldomero Oliva](#) [Javier García](#)

[Emre Güney](#) [Bioinformàtica](#) [Proteïnes](#) [Gens](#)

Comentaris

Nom:

Email:

Comentari:

0 comentaris

Cancelar

Publicar

