

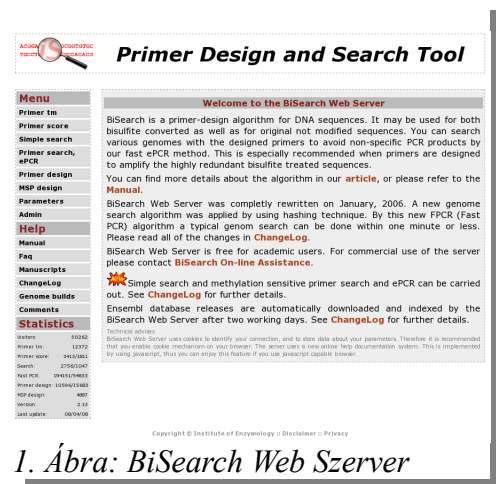
T61684 Sz. OTKA pályázat

Zárójelentés

Témavezető: Dr. Tusnády Gábor

A kutatás első évében, felkérésre összefoglalót készítettünk a humán ABC fehérjék membrán topológiájáról, kiegészítve az irodalmi ismereteket az általunk összeállított adatbázis statisztikus elemzésével. Az ABC családba tartozó fehérjék rendkívül fontos szerepet játszanak az emberi szervezet kifejlődésében, illetve működésében, valamint számos genetikailag öröklődő betegség, mint pl a cisztikus fibrózis, szintén az ABC transzporterekhez köthetők. Az eukarióta ABC transzporter fehérjék szerkezetéről nem állnak rendelkezésre nagy felbontású képek, ezért bioinformatikai módszerekkel - pl. az általunk kifejlesztett HMMTOP algoritmussal, illetve kísérleti és elméleti munkák kombinációjával - nyerhető információk ezen a területen különösen fontosak. Ebben a munkában valamennyi ismert humán ABC transzporter domén elrendeződésére adtunk igen megbízható becslést, és megmutattuk, az egyes alcsaládokra jellemző doménszerkezeteket is. Cikkünk a FEBS Letters-ben jelent meg.

Ebben az évben befejeztük továbbá a fehérje kutatások egy fontos segédeszközének, az általunk korábban kifejlesztett PCR primer-tervező szoftvernek egy jelentősen továbbfejlesztett, a korábbinál legalább két nagyságrenddel gyorsabban működő változatának kifejlesztését. A változtatás lényege, hogy az egyes genomiai szekvenciákból olyan index fájlokat készítünk, amelyek felhasználásával 16 nukleotidból álló oligomerek genomiai pozíciója a másodpercek tört része alatt megadható, akár a PCR primerek



1. Ábra: BiSearch Web Szerver

kitapadási tulajdonságát figyelembevevő mutációk megengedésével. A módszer egyedülálló továbbá abban a tekintetben is, hogy nemcsak módosítatlan genomiai szekvenciákon lehet az eljárás

segítségével elektronikus PCR-t végezni, hanem biszulfittal kezelt mintákon is, így a DNS metiláció közvetlenül vizsgálhatóvá vált. Az új algoritmussal működő szervert elhelyeztük a világhálón is (<http://bisearch.enzim.hu>, 1. Ábra). A szerver használhatóságát, illetve az általa nyújtott szolgáltatás fontosságát mutatja, hogy az elmúlt két év során több mint 50 ezer egyedi kutató használta, illetve az, hogy közel 25 ezer primer tervezésre és több mint 250 ezer (!) elektronikus PCR elvégzésére használták fel. A BiSearch eljárás továbbfejlesztéséről a BMC Bioinformatics-ban számoltunk be, valamint felkérésre írtunk egy összefoglaló könyv fejezetet is erről a témáról, ami a Methods in Molecular Biology-ban jelent meg.

TOPDB: Topology Data Bank of Transmembrane Proteins
Topology, Structure and Prediction.

Database Statistics

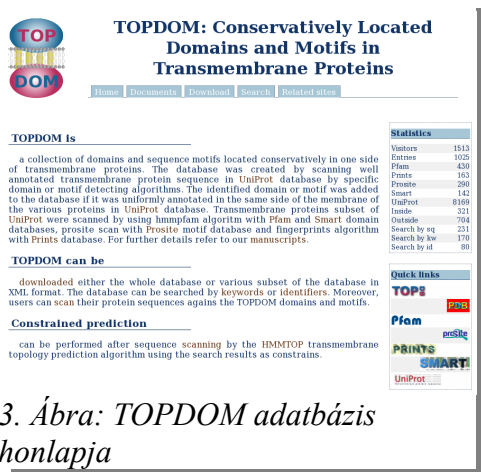
| | |
|-------------------------|----------|
| Database entries: | 1 |
| Revision date: | 09/20/07 |
| Enzymes: | 1497 |
| Topology data: | 2194 |
| Alpha helical proteins: | 1402 |
| Beta sheet proteins: | 42 |
| Published links: | 389 |
| TMHMM links: | 294 |
| UniProt links: | 1479 |
| Visitors: | 388 |

2. Ábra: TOPDB adatbázis honlapja

A kutatás második évében befejeztük a transzmembrán fehérjék topológiai adatbázisának összeállítását (TOPDB adatbázis). Ez az adatbázis az eddigi legátfogóbb topológiai adatbázis, amely nemcsak a már ismert szerkezetű transzmembrán fehérjék TMDet algoritmussal meghatározott topológiai adatait tartalmazza, hanem az irodalomban található különböző fizikai-kémiai, molekuláris biológiai módszerrel nyert topológiai adatokat is.

Ezeket az adatokat több száz cikk szisztematikus feldolgozásával nyertük ki. A kísérleti adatokat egységes és kereshető formában rögzítettük az adatbázisban ezáltal lehetővé vált, hogy ezeket az adatokat az elvégzett kísérletekben levő fehérjéken kívül más rokon és nem rokon transzmembrán fehérjék szerkezet és topológia becslésére is felhasználjuk. Az összegyűjtött topológiai adatokat felhasználva az adatbázisban levő valamennyi fehérje legvalószínűbb topológiáját is megadtuk a korábban általunk írt, de erre a feladatra – a pályázat munkatervével megfelelően - módosított HMMTOP eljárás segítségével. A HMMTOP eljárás módosítását eddig még külön nem közöltük, mivel további módosításokat

szeretnénk a módszer hatékonyságának növelésének érdekében, de a TOPDB adatbázist ezen módosítások nélkül nem tudtuk volna létrehozni. A TOPDB adatbázist az interneten az akadémiai kutatók számára hozzáférhetővé tettük (<http://topdb.enzim.hu>, 2. Ábra). Az adatbázis sikerét mutatja, hogy a megjelenését követő rövid időn belül több ezer egyedi látogatót regisztráltunk, illetve az egyedi és teljes adatbázis letöltések száma is igen magas. Ez az adatbázis szolgáltatja továbbá jelenleg a legnagyobb és legpontosabb topológia becsléshez használható ún. becslő halmazt is, így közvetett módon új topológia becslő eljárások kifejlesztését is segíti a TOPDB adatbázis. Az adatbázisról szóló közleményünket a Nucleic Acids Research folyóiratban közzétettük.



3. Ábra: TOPDOM adatbázis honlapja

A pályázat utolsó évében létrehoztunk még egy adatbázist, amely szintén a transzmembrán fehérjék szerkezetéről és topológiájáról nyújt értékes információkat, illetve az adatbázisban található adatokat direkt módon felhasználhatóvá tettük eddig még ismeretlen szerkezetű transzmembrán fehérjék topológia becslésére. Az adatbázis elkészítéséhez az a megfigyelés segített hozzá, hogy bizonyos szerkezeti és/vagy funkcionális doménok és szekvencia motívumok egyes transzmembrán fehérjékben mindig a membrán egy bizonyos, vagy a külső, vagy a belső oldalán találhatók meg. Ezen doménok azonosítására kidolgoztunk egy adatbányászati eljárást, amely az ún. UniProt adatbázisban levő transzmembrán fehérjék topológiai adatait felhasználva a különböző domén és szekvencia motívum adatbázisokban levő adatokhoz azok térbeli lokalizációját rendeli hozzá. Azok a domén és szekvencia motívumok kerültek be a újonnan létrehozott TOPDOM adatbázisba, amelyek térbeli lokalizációja konzervatív módon a membrán egyik vagy másik oldalán van. Az adatbázisba került adatok

közvetlenül felhasználhatók ismeretlen szerkezetű transzmembrán fehérjék topológia becslésére - a TOPDB adatbázisban létrehozásánál is használt - kényszerfeltételeket figyelembe vevő predikcióra („constrained prediction”) képes HMMTOP eljárás segítségével. A TOPDOM adatbázis szintén elérhető az interneten (<http://topdom.enzim.hu>, 3. Ábra). Az adatbázist leíró közleményünket a Bioinformatics folyóiratban közzeltük.

A transzmembrán fehérjék szerkezet becslésében elért eredményeink alapján jelen pályázat témavezetőjét felkérték a témáról szóló összefoglaló könyvfejezet megírására. A fejezet a Wiley kiadó „Wiley Book Series on Bioinformatics” sorozatában a „Protein Structure Prediction: Method and Algorithms” részben fog megjelenni.

A pályázatban közreműködő Simon István további 16 cikket publikált a jelen pályázat futamidejében, amelyek részben a jelen pályázat támogatásával készültek. Ezek közül 13 cikk a T049073 sz. OTKA pályázat zárójelentésében kerültek kifejtésre, illetve 3 a T072569 sz. OTKA pályázat első részjelentésében lesz található. Természetesen az itt felsorolt munkák elvégzéséhez, publikálásához az OTKA-n kívül más forrásokat is felhasználtunk de a támogatási források és a közlemények átfedés nélküli megfeleltetése lehetetlen egy dinamikusan működő kutatócsoport számára.