



**Magyar Agrár- és Élettudományi Egyetem
Budai Campus
Élelmiszertudományi és Technológiai intézet
Biomérnök alapképzési szak**

**Tejsavbaktérium eredetű béta-galaktozidáz enzim
tanulmányozása bioinformatikai eszközökkel**

Belső konzulens: **Dr. Bujna Erika**
 egyetemi docens
 Dr. Nguyen Duc Quang
 egyetemi tanár

Belső konzulens intézete/tanszéke:
 Élelmiszertudományi és Technológiai Intézet,
 Biomérnök és Erjedésipari Technológia Tanszék

Készítette: **Burján Alex**

Budapest

2024

Kutató munkám során 6 mikroorganizmus (*L. helveticus*, *L. acidophilus*, *L. casei*, *L. fermentum*, *L. plantarum* *L. reuteri*) által termelt béta-galaktozidáz enzimet hasonlítottam össze in silico módszerekkel. NCBI adatbázist használva az összes fent említett baktériumnak kerestem egy-egy kromoszómát. Átnézve a kromoszómán kódolt géneket összeírtam minden olyan gént, ami laktáz enzimet kódol. Összesen 12 gént találtam a 6 mikroorganizmus között. Ebből négy (*L. helveticus*, *L. acidophilus* *L. plantarum* *L. reuteri*) egy kisebb (*LacM*) és egy nagyobb (*LacL*) génen kódolja laktáz enzimet, a *L. fermentum* kromoszómáján csupán egy *LacM* található, míg az *L. casei* egy génen tárolja a teljes fehérjét. Ezt követően mindegyikhez hozzáadtam egy-egy ideiglenes nevét a következő módszerrel. Mindegyik gén nevének az első karaktere nagy L betű a második és a harmadik, pedig az adott baktérium nevének első két betűje, a negyedik és az ötödik betű minden esetben BG, mint béta-galaktozidáz. Utána folytatólagosan sorszámoztam az azonos kromoszómán található géneket. Pl. az *L. helveticus* előbb kódolt laktáz génje, *LheBG1*

Ezeknek a géneknek először összeírtam a konzervatív régióit, leggyakrabban a béta galaktozidáz kis lánc konzervatív régió volt tapasztalható a vizsgált fehérjék esetében.

Ezt követően többszörös szekvencia illesztést alkalmazva két filogenetikai fát készítettem a génekről. Az egyik fa csak a *LacL* és az *L. casei* géneket vette alapul, míg a másik csak a *LacM* géneket.

Az NCBI Phenogram adatbázis segítségével összeírtam a gének relatív elhelyezkedését a vizsgált kromoszómákon, majd azokat vizualizáltam.

A következő lépésben a Meme suite oldalon található eszközöket használtam fel, elsőként feltöltöttem az általam vizsgált gének szekvenciáját, és az oldalon használható xstreme eszköz segítségével megállapítottam, hogy a vizsgált enzimeknek várhatóan öt motívuma lehetséges. Ezt az információt tovább vittem az oldalon található MEME eszközbe, amellyel a fő motívumok elhelyezkedését ábrázolni tudtam. A legtöbb gén hasonló eredményt adott, a kiugró kivételek a *LacBG1* és a *LacBG2* voltak, amik már a filogenetikai vizsgálatok során is elkülönülést mutattak. Következésképpen az NCBI adatbázis open reading frame eszközét vettem használatba. Ennek segítségével azonosítani tudtam a vizsgált gének kódoló és nem kódoló régióit. Miután betápláltam ezeket az adatokat a Gene structure display szerverre ezeket ábrázoltam. Az volt megfigyelhető, hogy minden az általam vizsgált mikroorganizmusok közül a kódolt gének harmada kódoló régió, az egyetlen kivétel az *L. casei*

volt hiszen ezen baktérium esetében a laktáz enzim nem két külön részben kerül kódolásra a kromoszómán.

Ezt követően az enzimek fizikai-kémiai tulajdonságait vizsgáltam meg, Ehhez ProtParam-EXPASY internetes eszközt használtam. Az eszközt felhasználva össze tudtam írni a fehérjék molekulatömegét, izoelektromos pontját, negatív- és pozitív töltésű aminosavainak számát, instabilitási mutatóját, alifás mutatóját, valamint Gravy-mutatóját. Összehasonlítva az értékeket általánosan elmondható volt, hogy a LacM géneken kódolt fehérjék átlagosan 3500 Da tömegűek, míg a LacL 7000-8000 Da közötti tömeget érik el. Valamennyi a vizsgált fehérjék közül izoelektromos pontja alapján a savas tartományba esik. Az instabilitási mutató alapján a LheBG2, LacBG3, LreBG1 és a LreBG2-en kódolt fehérjék instabilisnak, míg a többi fehérje stabilnak tekinthető. A vizsgált enzimek közül, az összes hőstabilis és poláris molekula.

Végző lépésként a CELLO internete eszközbe, feltöltöttem a vizsgált enzimek szekvenciáját és az elemzést elindítva megkaptam a legvalószínűbb elhelyezkedésüket a sejten belül. Ez alapján az összes enzim intracellulárisnak mondható.