



Budapesti Műszaki és Gazdaságtudományi Egyetem
Vegyésszmérnöki és Biomérnöki Kar

A hasadó élesztőgombák *eng1* és *eng2* sejtosztódási génjeinek filogenetikai törzsfája

Tudományos Diákköri Dolgozat

Készítette: Vida Nóra

II. éves MSc biomérnök hallgató

Témavezetők:

Dr. Sveiczter Ákos *egyetemi docens*

Dr. Sipiczki Mátyás *tanszékvezető egyetemi tanár*

BME Alkalmazott Biotechnológia és Élelmiszer-tudományi Tanszék

DE TTK Genetikai és Alkalmazott Mikrobiológiai Tanszék

2011

Tartalomjegyzék

1. Bevezetés	3
2. Irodalmi áttekintés	5
2.1. Az eukarióta sejtciklus	5
2.2. Az eukarióta sejtciklus molekuláris szintű szabályozása	7
2.3. Dimorfizmus	8
2.4. A hasadó élesztő életciklusa	10
2.5. Citokinézis	11
2.6. A hasadó élesztő citokinézise.....	17
2.7. Filogenetikai fa	22
3. Célkitűzés	25
4. Módszerek	26
4.1. Genomikai adatgyűjtés és adatfeldolgozás... ..	26
4.2. Fehérje szerkezet vizsgálat	30
4.3. Többszörös szekvencia illesztés	32
4.4. Törzsfa készítés	32
5. Eredmények.....	34
5.1. Szekvencia illesztés.....	34
5.2. Fehérje domének azonosítása.....	38
5.3. Többszörös illesztés és faszerkesztés.....	39
6. Összefoglalás.....	45
7. Irodalomjegyzék.....	47
8. Függelék	50

1. Bevezetés

Minden élő organizmus elemi építőköve a sejt, mely önmagában is élőknek tekinthető, hiszen anyagcserére és szaporodásra is képes. A sejtnövekedés és ezt követően a sejtosztódás alapvető fontosságú a sejtek működéséhez és fejlődéséhez, ezért a sejt magas szinten szabályozott események sorozatával, a sejtciklus segítségével hajtja mindezt végre. A sejtosztódás a sejtek szaporodásának folyamata, mely eredményeként önmagukkal többé-kevésbé azonos utódsejteket hoznak létre. A sejtek élete a születésüktől az osztódásukig tart. Az eukarióta organizmusok sejtciklusa evolúciósan nagyon konzervált, így azok alapvető komponensei már 1,5 milliárd éve megjelenhettek és kialakulhattak.

A kutatók gyakran egyszerűbb felépítésű eukarióta mikroorganizmusokat választanak kísérleteik alanyául, amelyek működése nagyban hasonlít a magasabb rendű eukarióta sejtekéhez, azonban velük ellentétben könnyebb és kevésbé költséges a kutatás. Az élesztők a huszadik század első felében a sejtbiológia számos területének modell organizmusaivá váltak. A hasadó élesztőként is ismert *Schizosaccharomyces pombe* az ötvenes évek vége óta lett egyre kedveltebb modellorganizmusa a citológiai és genetikai kutatásoknak, többek között kedvező fiziológiai tulajdonságai miatt (Kim, 2010). Ezt az egysejtű gombát, más néven hasadó élesztőt, lekerekedett végű, henger alakú sejtek jellemzik, a születéskori sejt hossza körülbelül 7-8 μm , átmérője pedig 3-4 μm . Sejtciklusa során a *S. pombe* hosszirányban folyamatosan nő, átmérője azonban nem változik, osztódást követően pedig két közel azonos méretű leánysejtet hoz létre, tehát osztódása szimmetrikusnak mondható.

A sejtosztódás vagy más néven citokinézis az eukarióta sejtciklus utolsó lépése, amely során egy anyasejtből két leánysejt jön létre. A *S. pombe* jellegzetessége, hogy a citokinézis nem következik be a mitózis végén, hanem csak később valósul meg. Ennek oka, hogy szemben pl. az állati sejtekkel, a gombák sejtfallal rendelkeznek, tehát nem tudnak egyszerűen a plazmamembrán befűzésével elosztódni a mitózis végén, hanem meg kell szüntemizni az új végekhez tartozó sejtfallat. Ezt a folyamatot nevezzük szeptumképzésnek; a szeptum pedig egy különböző poliszacharid rétegekből álló struktúra. Ezen poliszacharid rétegek közül a középsőnek, amelyet elsődleges szeptumnak is neveznek, el kell hidrolizálnia sejtosztódás közben. Ez az elsődleges szeptum *Candida albicans* esetében pl. kitint is tartalmaz, míg a *Schizosaccharomyces*

nemzetség esetében β -1,6 kötések is tartalmazó β -1,3-glükánból és β -1,6-glükánból áll. Különböző fajok speciális enzimeket termelnek, hogy ezen struktúrákat le tudják bontani, így pl. kitinázt és β -endoglükánázt. *S. pombe* esetében azonosították az *eng1* és *eng2* géneket, amelyek által kódolt fehérjék endo- β -1,3-glükánáz aktivitással rendelkeznek. Az *eng1* felelős az elsődleges szeptum hasításáért (Martín-Cuadrado et al., 2003), az *eng2*-nek pedig a sporulációban van szerepe (del Dedo et al., 2009).

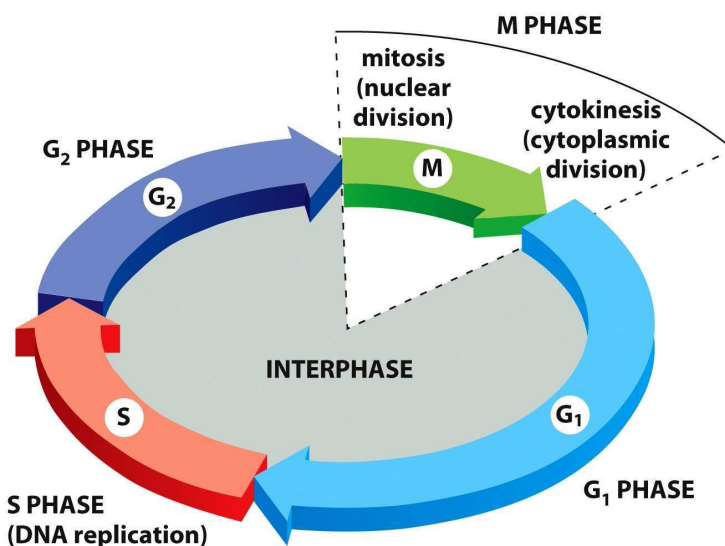
A gombák a növények és az állatok mellett önálló evolúciós irányvonalat képviselnek, mind a növényekre, mind pedig az állatokra jellemző tulajdonságaik is vannak. Az Ascomycota vagy tömlősgombák tagozatába tartozik a dimorfizmust mutató *Schizosaccharomyces japonicus*, mely egysejtű élesztő és többsejtű fonalas alakban is képes növekedni (Enczi et al., 2007). Filogenetikai értelemben az látszik valószínűnek, hogy a fonal az ősi és az élesztő az újabb morfológiai megjelenés. A hasadó élesztő sejtosztódásában szerepet játszó gének (*eng1* és *eng2*) szekvencia-analízisével ismert fajokkal való rokonsági kapcsolatokat és a közöttük lévő evolúciós távolságokat szeretném jobban megismerni, többek között a *Candida albicans* humán patogén dimorf gombát is érintve.

A hasadó élesztő jelentőségét növeli az is, hogy 2002 óta teljes genomja ismert (Wood et al., 2002). Emellett érdemes megemlíteni, hogy a 2001-es orvosi Nobel-díjat az eukarióta sejtciklussal foglalkozó három kutatónak (Leland Hartwell, Tim Hunt, Paul Nurse) ítelték meg, akik eredményeikkel új lehetőségeket teremtettek a rák kezelésében (Nasmyth, 2001). A három kutató közül *Paul Nurse* szinte egész életében a hasadó élesztő sejtciklusával foglalkozott, és ezen kutatásaival érte el pályájának legfontosabb eredményeit.

2. Irodalmi áttekintés

2. 1. Az eukarióta sejtciklus

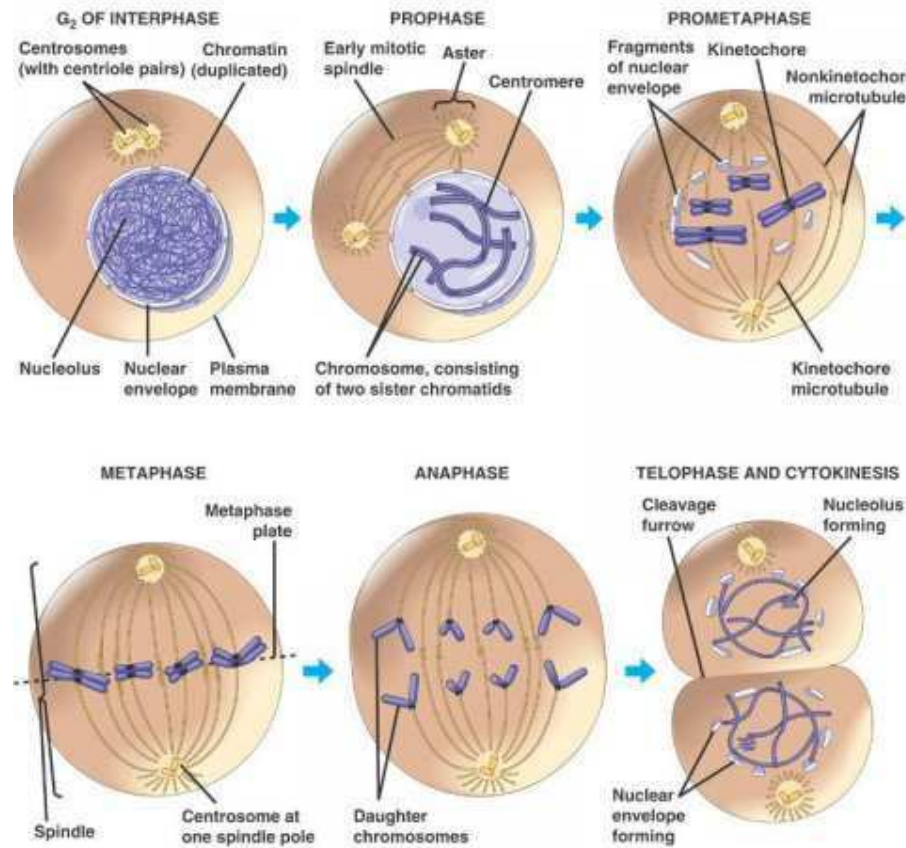
A sejtciklus legfontosabb eseményei a DNS precíz másolása, azaz replikálása, majd ezt követően a másolatok pontos szétosztása, hogy genetikailag azonos leánysejtek keletkezhesse. A legtöbb sejt a ciklusa során megduplázza a tömegét és a citoplazmás organelleumait is.



1. ábra. Az eukarióta sejtciklus (Alberts et al., 2009)

Az eukarióta sejtciklus négy fázisra osztható (1. ábra). Két jelentős esemény zajlik a sejtciklus végén: a sejtmag osztódása, a mitózis, azt követően pedig a sejt kettéosztódása, vagyis a citokinézis. E két esemény együttesen alkotja az M-fázist. Interfázisnak nevezzük a két M-fázis közti szakaszt, amely további három periódusra tagolható, a G₁-, az S- és a G₂-fázisokra. Az S-fázis alatt történik a nukleáris DNS replikációja, ami elengedhetetlen a későbbi sejtosztódáshoz. Azonban az S-fázis az interfázisnak csak egy kis része, így az S és M-fázisok közt szünetek, angolul gap-ek vannak. A G₁-fázis a mitózistól az S-fázisig tart, a G₂-fázis pedig az S-fázistól a mitózisig. A szünetek alatt a sejt mind a külső, mind pedig a belső környezetét ellenőrzi, hogy azok megfelelőek legyenek a ciklus további eseményei számára. A G₁- és G₂-fázisok együttesen teszik lehetővé, hogy a sejtnek elegendő ideje legyen a növekedésre

és a citoplazmás organelumok számának megduplázására. A mitózis az M-fázis során játszódik le, ekkor a sejt egyenlően elosztja a leánykromatidákat a két leánysejt magjaiba (Morgan, 2007).



2. ábra. A mitózis fázisai

(http://royaleb.files.wordpress.com/2009/04/mitosis_phases1.jpg)

A mitózis a citológiailag megfigyelhető események alapján öt egymást követő szakaszra bontható (2. ábra.). A mitózis első szakasza, vagyis a *profázis* alatt a kromoszómák kondenzálódnak és a testvér kromatidákat kohezinek tartják össze. A sejtmagmembrán a *prometáfázis* során bomlik le, kialakul a mitózisos orsó, ami a kinetokórnál kötődik a kromoszómákhoz. A metafázisos síkra a kromoszómák a *metafázis* alatt állnak fel, az orsó húzófonalai a kromoszómákhoz tapadnak. A mitózis *anafázisa* közben a testvér kromatidák elválnak egymástól, és leány kromoszómává válnak. A fajok többségénél az anafázis két részből tevődik össze, az *anafázis A* során a

kinetokór mikrotubulusok rövidülnek, ezért a leánykromatidák a centroszómák felé vándorolnak, az *anafázis B* alatt pedig az interpoláris mikrotubulusok hosszabbodása következtében a centroszómák egymáshoz képest távolodnak. Ezután újra összeáll a nukleáris membrán, és két leány sejtmag keletkezik, ezt a szakaszt *telofázisnak* nevezik. Az M-fázis utolsó szakaszában a sejt citoplazmája is kettéosztódik. Fontos megemlíteni, hogy a gombák ún. zárt mitózissal rendelkeznek, amely során a magmembrán nem oldódik fel, hanem a sejt mitózisa a nukleuszon belül játszódik le.

2.2. Az eukarióta sejtciklus molekuláris szintű szabályzása

Egy központi kontrollrendszer indítja el a sejtciklus egyes folyamatait, amely arról gondoskodik, hogy a helyes működése közben a genomi DNS precízen megduplázódjon, majd szegregálódjon. Mindezt regulátor fehérjék komplex hálózata biztosítja, ami maga a kontrollrendszer.

Három ellenőrzési pont van a sejtciklus alatt. Az első ellenőrzési pont a G1/S fázisok határán van: ekkor a sejt ellenőrzi, hogy a környezet kedvező-e a ciklusba lépéshez, azaz a replikáció iniciálásához és hogy az előző mitózis rendben végbement-e. Ha az extracelluláris körülmények nem megfelelőek, akkor a sejt késleltetni tudja a G1-fázis előrehaladását és esetleg beléphet egy speciális nyugalmi állapotba, melyet fejlettebb élőlényeknél G0-fázisnak neveznek. Ezt a pontot, ami a G1-fázis vége felé található, élesztőkben start pontnak (Forsburg and Nurse, 1991), magasabb rendű szervezetekben restrikciós pontnak nevezzük (Zetterberg et al., 1995). A következő ellenőrzési pont a G2-fázisban fejt ki hatását, mely biztosítja, hogy a sejt ne lépjen a mitózisba, amíg az esetlegesen sérült, károsodott DNS-t ki nem javítja és a DNS replikáció nem teljes (Maller, 1994). A harmadik ellenőrzési pont a mitózisban, a metafázis/anafázis határán hat, amely ellenőrzi, hogy minden leánykromatida megfelelően kapcsolódott-e a kinetokór mikrotubulusokhoz (Murray, 1994).

A sejtciklus kontrollrendszer fehérjéi feltehetően már több, mint egy milliárd évvel ezelőtt jelentek meg, és az evolúció során konzervált formában maradtak meg. A sejtciklus kontrollrendszer a proteinek két családján alapul, a ciklin dependens protein kinázokon (Cdk) és a ciklineken. A kontrollrendszer irányítja a sejtciklus mechanizmusát azáltal, hogy ciklikusan aktiválja és inaktíválja a fehérjéket és komplexeiket, melyek elindítják és szabályozzák a DNS replikációt, a mitózist és a

citokinézist (Morgan, 2007). Az egyik legáltalánosabb út, ahogyan a sejt aktiválja, illetve inaktiválja fehérjéit, a foszforiláció és az azt követő defoszforilációs lépések bekövetkezése. A foszforilációs reakciók, amik a sejtciklust szabályozzák, egy speciális protein kináz családnak köszönhetők. Ezek a fehérjék csak a ciklus megfelelő időpontjában aktiválódnak, majd gyorsan ismét inaktívvá válnak. A kinázok aktiválása a ciklus megfelelő időpontjában egy másik csoport fehérjének a feladata a kontrollrendszerben, a ciklineké. A ciklineknek nincs enzimatis aktivitásuk, de kapcsolódniuk kell a kinázokhoz, még mielőtt azok aktiválttá válnának, ezért ezeket a kinázokat hívják ciklin dependens protein kinázoknak vagy röviden Cdk-nak. Tehát a két fehérjealegységből álló komplex faktor katalitikus komponense egy specifikus protein kináz molekula, amely akkor nyerheti el aktivitását, amikor aktivátorával, a ciklin nevű regulátormolekulával heterodimert alkothat (Sveiczter et al., 2004).

A különböző lépéseket a ciklus során különböző ciklin-Cdk komplexek indítják el. Azt a ciklint, ami a G2-fázisban hat és az M-fázisba lépést idézi elő, M ciklinnek hívják, és azt az aktív formát, melyet a megfelelő Cdk-val alkot, MPF-nek (M-phase promoting factor) nevezik. A DNS replikáció előtt egy másik Cdk hozzákapcsolódik az S fázisú ciklinhez, hogy ezáltal kialakítsák a SPF-eket (S-phase promoting factor), illetve hogy előidézzék az S-fázist. A különböző komplexek különböző átalakulásokért (változásokért) felelősek a sejtben. A Cdk-k aktivitását szintén a foszforiláció és defoszforiláció szabályozza (Alberts et al., 2009).

A mitózisba lépést egy citoplazmás regulátor, az MPF szabályozza, mely univerzális az eukariótákban. Az MPF aktivitás oszcillációja szabályozza a sejtciklust. Az MPF tisztítása során kiderült, hogy az egy szerin és treonin oldalláncokat foszforilező protein-kináz, melynek két alegysége van, a Cdc2 és a ciklin-B. Az aktív MPF indukálja a mitózis folyamatait, hatására a kromoszómák kondenzálódnak, a nukleáris membrán lebomlik, a citoszkeleton átrendeződik, és kialakul a magorsó (Alberts et al., 2009).

2.3. Dimorfizmus a gombák világában

A gombákra kétféle növekedési mód jellemző: az egysejtű élesztő és a többsejtű fonalas (hifa vagy pszeudohifa) alak. Ezek közül az egyes gombák általában jellemzően csak az egyik formában fordulnak elő. A di- és polimorf gombák azonban a környezeti

hatásokhoz igazodva képesek váltani az élesztőkhöz hasonlító egysejtű és a fonalas növekedési formák között. Többek között a nem-patogén, szaprofita *Schizosaccharomyces japonicus* gomba is képes egysejtű élesztőként és többsejtű, valódi micéliumokat képző organizmusként is jelen lenni környezetében. Ezen tulajdonsága miatt vált fontos modell organizmusává a patogenitást célzó kutatásoknak, a di- és polimorfizmus vizsgálatán keresztül. A növénypatogén *Mucor*, *Rhizopus*, *Ustilago* nemzetségek, illetve a humán patogén *Candida albicans* többsejtű hifaalakban virulens míg a legtöbb, szisztémás mikózist okozó humánpatogén gombafaj, pl. a *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Paracoccidioides brasiliensis* egysejtű élesztőformában okozza a megbetegedést.

Számos humán és növényi patogén gomba esetében a képesség, hogy tudjon változtatni növekedési módján, összefüggésben áll az adott patogén fertőzőképességével, mivel a fertőzés mechanizmusa közben gyakran szükséges a morfológiai átmenet. A valódi hifák szeptumok által elkülönített sejtekből épünek fel, a hifasejtek között van citoplazmás áramlás, és így kommunikáció is (Sipiczki et al., 1998). A *Saccharomyces cerevisiae* is fontos organizmusa a dimorfizmus kutatásoknak, hiszen képes sarjadzó élesztő sejtek és pszeudohifák képzésére is, ellenben valódi hifák képzésére nem. A pszeudohifa láncszerűen összekapcsolódott, megnyúlt sejtekből áll, azonban a valódi hifákkal ellentétben a szeptumoknál összeszorított, és a szeptumoknál kezdődő elágazásokat tartalmaznak, amely elágazások végén található sejtek kisebbek a többi sejtnél. Az *S. cerevisiae*-vel ellentétben sok di- és polimorf gomba képez valódi hifát vagy valódi hifát és pszeudohifát is, többek között ilyen a *Candida albicans*, *Ustilago maydis*, *Cryptococcus neoformans*, és a *S. japonicus* is. A *S. japonicus* folyékony környezetben élesztőszerű sejteket formál, szilárd szubsztrátok esetében azonban képes valódi hifákat vagy pszeudohifákat alkotni. A *S. japonicus* valódi hifái számos patogén fonalas organizmushoz hasonlóan invazív természetűek. Az élesztőből hifák formálásává átváltó folyamat fokozatosan megy végbe, amelyhez szükséges a sejteknek a bipoláris növekedési formából unipolárisrá változnia, emellett további feltétele a sejtek megnyúlása és a sejtelválás megszűnése. A sejteknek a bipoláris növekedési formából unipolárisrá való átalakulása összefüggésben áll a sejt egyik végén létrejövő nagyméretű vakuólum megjelenésével, amely gátolja a sejt hozzá közel eső végének további növekedését (Enczi et al., 2007). A gombák addig képesek többsejtű fonalas alakban nőni, amíg szilárd szubsztrát rendelkezésükre áll, azonban a környezet

tápanyagtartalmában történő hirtelen változásra újra képesek egysejtű élesztő alakban növekedni (Sipiczki et al., 1998).

Az élesztő alakból pszeudohifává, illetve valódi hifává alakulás szabályozásában szerepet játszó gének számos kutatás középpontjába kerültek, hiszen fontos szerepük lehet újszerű gombaellenes szerek kifejlesztésében (Enczi et al., 2007).

2.4. A hasadó élesztő életciklusa

A *Schizosaccharomyces* nemzetségbe négy fajt tartozik: a *S. pombe*, a *S. japonicus*, a *S. octosporus* és a *S. cryophilus*. A nemzetség tagjait más néven hasadó élesztőknek is nevezik, melyek a tömlősgombák (*Ascomycota*) tagozatában az *Archiascomycetes* osztályba tartoznak (Rhind et al., 2011).

A *Schizosaccharomyces pombe* genomja hozzávetőlegesen 14,1 millió bázispárból áll, mely körülbelül 4970 gént tartalmaz. A hasadó élesztőt 1893-ban Lindner izolálta először egy kelet-afrikai sörből, neve is innen ered, a *pombe* szuahéli nyelven sört jelent. Modellorganizmusként először Urs Leupold használta az 1950-es években, örökléstani kutatásai alanyául. Előfordulásuk erjedő növényi anyagokban, mustban gyakori. Ivaros szaporodásuk a sejtek konjugációjával, majd askospórák képzésével történik. A *S. pombe* haplonta faj, melynek diploid állapota létezik ugyan, de nem stabil, hanem meiózissal haploid sejtekké alakul. A hasadó élesztő haploid sejtjei kétféle, h^+ vagy h^- párosodási típussal rendelkeznek, és lehetnek heterotalliások vagy homotalliások (h^{90}). Az ivaros szaporodás során két ellentétes (h^+ és h^-) párosodási típusú sejt konjugál egymással és így diploid zigóta jön létre, mert a plazmogámia és a kariogámia szinte egyidejűleg következik be. A diploid zigóta az esetek nagy részében rögtön askusszá alakul, amelyben meiózissal 4 askospóra képződik. A haploid askospórák kicsírázásával vegetatív sejtek keletkeznek, amelyek ivartalanul osztódva tenyészeteket hoznak létre (Sipiczki, 1995).

Kedvező körülmények közt vegetatív módon a haploid sejtek hasadással szaporodnak. Ivartalan ciklusa során a *S. pombe* pálcika alakú egysejtű eukarióta, melynek átmérője 3-4 μm . A henger alakú sejtek kizárólag csak hosszirányban nőnek a végeiken, a sejt közepén nincs növekedés (MacNeill and Fantes, 1995). A sejtek átlagos hossza születéskor 7-8 μm , osztódáskor pedig 13-14 μm .

A nemzetség további tagjai közül a *S. japonicus* és a *S. octosporus* életciklusuk során nyolc aszkospóra képzésére képesek (Sipiczki, 2000). A *S. japonicus* dimorf szervezet lévén képes mind élesztősejtek, mind pedig hifák képzésére. Az egy sejtű alakból hifákká való átmenet a környezet által szabályozott, csak szilárd táptalajon és tápanyag gradiens esetében alakul ki. A *S. cryophilus*, a legutóbb leírt hasadó élesztő, morfológiáját tekintve leginkább a *S. octosporus*-hoz hasonló, de a *S. cryophilus* esetében a sejtek átlagos mérete kisebb, mind hosszukban ($5,25 \pm 0,20 \mu\text{m}$ összehasonlítva $7,00 \pm 0,34 \mu\text{m}$), mind pedig átmérőjükben ($4,85 \pm 0,10 \mu\text{m}$ összehasonlítva $5,68 \pm 0,09 \mu\text{m}$). A *S. cryophilus*-nak az optimális növekedési hőmérséklete alacsonyabb, mint a többi hasadó élesztő fajnak, a nevét is innen kapta (Helston et al., 2010).

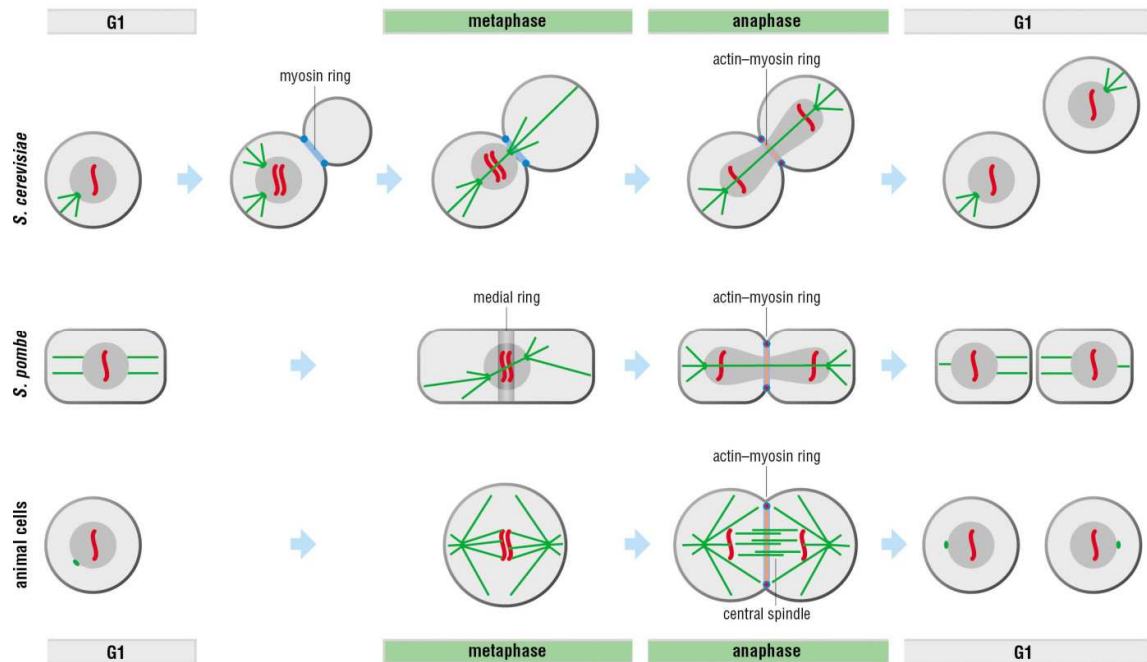
2.5. Citokinézis

A sejtosztódás utolsó lépését nevezik citokinézisnek. A mitózis befejeztével a két új sejtmag és más sejtorganellumok elkülönülnek két új leánysejtbe, amelyek mindegyikébe kerül egy sejtmag, egy centroszóma, és nagyjából egyenlő részben osztoznak a citoplazmás makromolekulákon és orgnillumokon. Általában az anafázisban kezdődik meg a folyamat, de a két leánysejt nem válik el, míg a nukleusz véglegesen ki nem alakul a telofázisban (Alberts et al., 2009).

A citokinézis eltérő módon jelenhet meg különböző organizmusokban (3. ábra). A pékélesztő osztódása közben egy sarjat növeszt, mely a sejtciklus közben növekszik, majd a mitózis után leválik az anyasejtről. A gombák és a növények a mitózist követően új sejtfalat hoznak létre, a szeptumot a sejt közepén.

A citokinézis a legtöbb eukariótában, kivéve a növényekben egy centripetális mechanizmus eredménye, amikor is a sejtmembrán befelé húzódik a hasadási oldalon és egy hasadási barázda alakul ki, amely aztán kettéválasztja az anyasejtet. Ez a hasadási mechanizmus a kontraktilis gyűrű irányításával megy végbe, amely a sejtet körülfogva a sejtkéreg (kortex) alatt jelenik meg, és kettéválasztja a szegregált kromoszómák tengelyeit. Állati sejtek esetében a kontraktilis gyűrű anafázis utáni fokozatos összehúzódása húzza befelé a membránt és eredményezi a hasadási barázda megjelenését. Gomba sejtek esetében is egy, az állati sejtekben megfigyelhető aktin-

miozin alapú gyűrűhöz hasonló struktúra működik, melynek összehúzódását az új sejtfal összeszerelődése követ (Bathe and Chang, 2009).

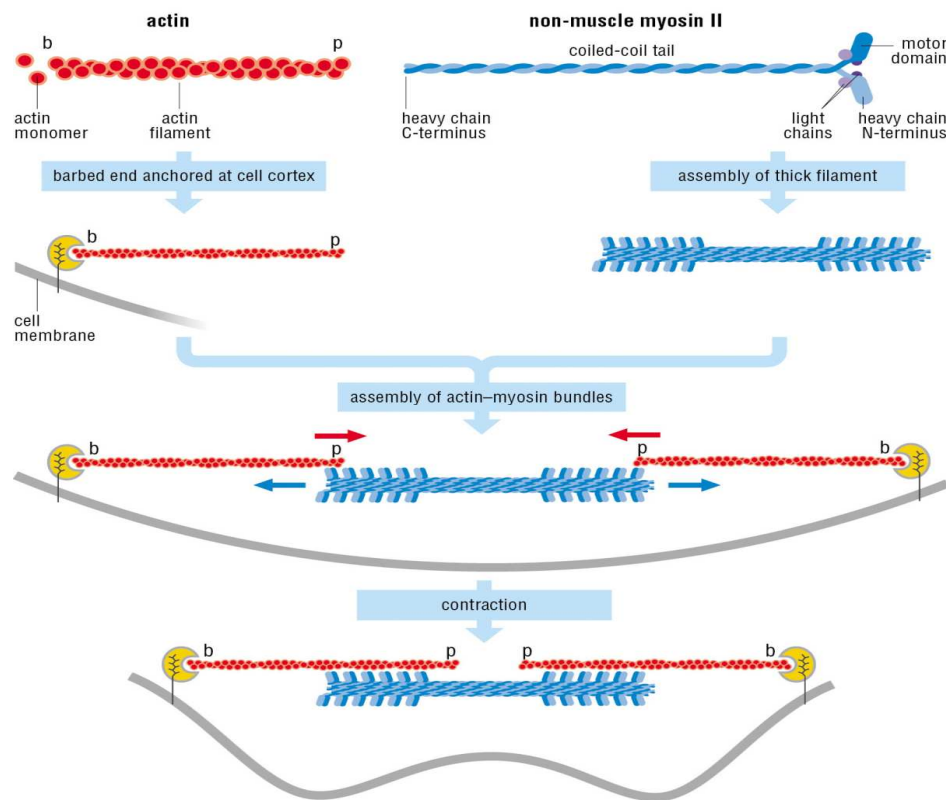


3. ábra: A sejtosztódási sík elhelyezkedése (Morgan, 2007)

A gyűrűt aktin szálak és miozin II motorfehérjék alkotják, amelyek kontraktilis köteget alkotnak együtt és összehúzódásuk eredményezi a gyűrű összezsugorodását és a membrán befelé húzódását. Az aktin egy hosszú láncokká polimerizálódó citoskeletonális fehérje, amely láncokat mikrofilamentumoknak vagy más néven F-aktinnak neveznek, ezen láncok két kötegből álló hélixet formálnak, ahogyan az a 4. ábrán látszik.

A mikrotubulusokhoz hasonlóan az aktin szálak is poláris jellegűek. Megkülönböztetünk egy horgas véget (barbed end), amely analóg a mikrotubulusok plusz végével, ez a vég gyorsabban nő, mint az úgynevezett hegyes vég (pointed end), amely analóg a mikrotubulusok mínusz végével. Azonban az aktin filamentumok vékonyabbak és flexibilisebbek, mint a mikrotubulusok, ezért sokféle struktúrába képesek rendeződni, úgy mint például a nagyon rendezett lineáris kötegekben fekvő vázizmoktól kezdve a különféle kereszttekötött, két- és háromdimenziós hálózatokig, amelyek a sejt alakjának megtartásában és mozgékonyágában segítenek. Az aktin

filamentumok a miozin motorfehérjék pályájaként szolgálnak (ugyanis azon mozogni képes), amelyek általában az aktin szál horgas vége felé utaznak.



4. ábra: Az aktin-miozin gyűrű (Morgan, 2007)

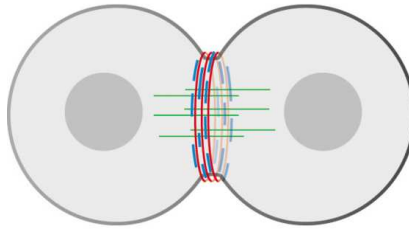
A legtöbb összehúzóási folyamathoz, beleértve a citokinézist is, az erőt valamilyen miozin II. motor fehérje mozgása eredményezi az aktin szál mentén. A miozin II több alegységből épül fel, többek között tartalmaz két szorosan összefüggő nagy alegységet is, az úgynevezett nehéz láncokat. A nehéz láncoknak a hosszú karboxi-terminális farok doménje egy egymás köré tekeredett (coiled-coil) formába alakul, ahogyan az a 4. ábrán is látható. Ezeknek a fehérjéknek az amino-terminális végén pedig globuláris fej-domének találhatóak, amelyek ATP-dependens motor aktivitással rendelkeznek. Az összehúzóási mechanizmus a miozin II. filamentumok bipoláris tulajdonságán alapul: a motor domének mindkét végen ellentétes irányba orientálódnak. Ennek eredményeként a miozin II. filamentumok képesek az aktin szálak hegyes végét befelé húzni az ellentétes irányból. Hogyha az aktin filamentumok horgas vége lehorgonyzott például a sejtkeregben, akkor a két lehorgonyzott pont

egymáshoz közelebb fog kerülni, amikor az aktin-miozin kapcsolat összehúzódik. A kontraktilis gyűrűben jelen lévő miozin II fehérje egy különleges, nem-izom eredetű vagy más néven citoplazmás miozin II. Mindkét nem-izom eredetű miozin II. motor domén egy-egy könnyű láncsal kapcsolódik össze (Sipiczki, 2007).

Állati sejtek esetében az aktin-miozin gyűrű összehúzódását elsősorban a szabályozó miozin könnyű lánc (regulatory myosin light chain - RMLC) foszforilációjának változása szabályozza. A nem izom eredetű miozin II. néhány szerin maradvány foszforilációja által aktivált, közel az RMLC amino végéhez. Az RMLC foszforilációja nem csak az ATP-dependens miozin II. motor aktivitásához szükséges, hanem elősegíti a miozin II. bipoláris filamentumokká való rendeződését is, amely szükséges az összehúzódó aktin-miozin kötegek kialakulásához is. Élesztőben nem figyelhető meg az RMLC foszforilációja, esetükben az Rng3 fehérje kapcsolódik és aktiválja a miozin II-t a citokinézis alatt (Morgan, 2007).

Az aktin filamentumok megjelenése a hasadási oldalon kialakulhat akár a már máshol megjelent korábbi filamentumok felépüléséből vagy a lokálisan kialakuló új filamentumból is. Mindkét lehetőség megvalósulhat, de az utóbbi nagyobb jelentőségűnek bizonyul. A hasadási oldalon az aktin filamentumok megjelenésében nagy szerepe van a fehérjék egy kis családjának, a forminoknak. Forminok segítségével valósul meg a szolubilis aktin monomerek összetömörülése egy új aktin szállá. A formin kölcsönhatásba lép az aktin filament horgas végével, miközben elősegíti az aktin monomerek bekapcsolódását ugyanezen a végen. Az aktin monomerek bekapcsolódását az újonnan szintetizált formin-sapkával ellátott horgas végű szálla egy profilin nevű fehérje segíti, amely az aktin monomerekhez kapcsolódik (Morgan, 2007).

A kontraktilis gyűrűben az aktin és miozin II szálla, a vázizmokban megfigyelhető jól szervezett struktúrával ellentétben, kevésbé rendezettek, nagyjából függőlegesen helyezkednek el az orsó tengelyére (5. ábra).

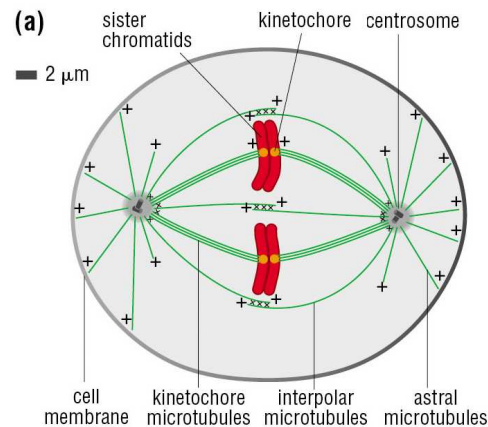


5. ábra: Az aktin-miozin kötegek elhelyezkedése a kontraktilis gyűrűben (Morgan, 2007)

A citokinezis során a gyűrű fokozatos összehúzódása közben, ez a dinamikus aktin-miozin hálózat szétszerelődik és újra rendeződik, ezáltal a fonál vastagsága nem változik szignifikánsan. Az aktin eltávolítása a csökkenő kontraktilis gyűrűről részben az aktin-destabilizáló fehérje, a cofilin működésén alapul.

A membrán és a sejtfal elmozdulása az aktin-miozin gyűrű megjelenésével párhuzamosan történik, ami azt jelenti, hogy a membrán növekedésének mértéke illeszkedik a kontraktilis gyűrű befelé húzódásának mértékéhez. Annak érdekében, hogy minden leánysejtbe egy genomi másolat kerüljön, a hasadási sík a két szegregált kromoszóma szerelvény között alakul ki. Élesztőkben a hasadási sík már a mitózis előtt meghatározott. A sejt a mitózis orsó segítségével helyezi el a szegregált kromoszómákat az ellentétes oldalakra (6. ábra). Pékélesztő esetében a sarj a késői G1-fázisban jelenik meg és növekszik a sejtciklus alatt, majd az anafázisban a mitózis orsó a sarj nyakában helyezkedik el, ezáltal helyezi el a kromoszómákat az egyes utódokban. Hasadó élesztőben a mitózis előtt, a hasadási oldalon a fehérjék gyűrű formában halmozódnak fel, ezáltal megjelölve a hasadás helyét a sejt közepén. A mitózis orsó pólusai pedig az ellentétes oldalán vannak ennek a gyűrűnek. Állati sejtekben a hasadási hely nincs meghatározva az anafázisig, amikor is a mitózis orsó küld egy jelet a sejtkeregnek, hogy indulhat a hasadási barázda képződése a szegregált kromoszómák között. A jel keletkezésének mechanizmusa és molekuláris alapja azonban nem tisztázott és valószínűleg eltérő az egyes sejtípusokban. Néhány sejtben például az asztrális mikrotubulusok küldik a jelet az orsó pólusából, más esetekben a hasadási jel a sejtkereghez a centrális orsóból (central spindle) jön egy köteg antiparallel mikrotubulus által, amik részben az anafázisos orsó interpoláris mikrotubulusaiból származnak (Burgess and F., 2005). A centrális orsó számos regulátor fehérjét tartalmaz, amelyek feltehetően többféle (összetett) feladatokat látnak el a citokinezis során: lehetséges, hogy nem csak a hasadási hely meghatározásában segítenek, hanem a

barázda befelé húzódásának irányításában is. Ahogyan a hasadási barázda mélyül, a kontraktilis gyűrű végül találkozik a centrális orsóval, amely összetömörül, és az így kialakult struktúrát midbodynak nevezzük, amely antiparallel mikrotubulusokat és vastag fehérje mátrixot tartalmaz (Morgan, 2007).



6. ábra: A mikrotubulusok szerepe a sejtosztódásban (Morgan, 2007)

A citokinézis nem történhet meg addig, amíg a kromoszómák nem szegregálódtak. A sejtosztódás utolsó lépése olyan mechanizmusok által iniciált, amelyek a mitózis befejezésével állhatnak közeli kapcsolatban. Úgy tűnik, hogy a citokinézis indításában jelentős szerepet tölt be az anafázisban a Cdk szubsztrátok defoszforilációja. Általában a citokinézis nem indul el addig, amíg a Cdk1 inaktíválódás meg nem történik (Morgan, 2007).

Pékélesztő esetében a sarj előkészítése a késői G1-fázisban kezdődik, amikor a szeptinek és más fehérjék jelennek meg széles gyűrű alakban a sarj keletkezési helyén. Szeptin függő miozin II. gyűrű alakul ki a sejt azon oldalán, ahol majd a sarj fog megjelenni. A citokinezis ezen korai lépései a G1/S-Cdk aktivitásának köszönhető, amely a G1-fázis végén növekszik (Alberts et al., 2009).

Hasadó élesztő esetében a hasadási hely a korai mitózis során alakul ki a sejt középső vonalán, ezzel egyidejűleg a Mid1 nevű fehérje is megjelenik a sejt közepén. Ennek a középső gyűrűnek az elhelyezkedését pedig a sejtmag helyzete határozza meg, amely általában a sejt közepén található (Bathe and Chang, 2009). Mitózis során számos, a kontraktilis gyűrű kialakulásához szükséges komponens, többek között a miozin II, az aktin, a formin Cdc12 és az Rng2 fehérje - amely az aktin filamentumok

helyes elrendeződéséhez szükséges a szeptumban – gyúlik össze a középső gyűrű közelében. Az anafázis után ezek alkotják majd a kontraktilis gyűrűt.

Általában eukariótákban a mitózis orsó a sejt közepén helyezkedik el, így a keletkező leánysejtek közel azonos méretűek lesznek, azonban előfordulnak olyan esetek is, amikor az orsó nem középen alakul ki, hanem a sejt egyik végéhez közel helyezkedik el, ennek következtében aszimmetrikus sejtosztódás valósul meg. Ilyen esetekben a leánysejtek nem csak méretükben nem azonosak, de az örökölt molekuláris összetételükben is különböznek és általában eltérő sejt típusokká fejlődnek.

Növényekben a citokinézis az állati sejtektől eltérő módon játszódik le, a két leánysejt nem a kontraktilis gyűrű hatására válik ketté a felszínen, többek között azért, mert a növényi sejteket egy sejtfal veszi körül. Az új sejtfal a *telofázis* elején kezd kialakulni a citoplazmában a két szegregált kromoszóma szerelvény között. Ezt a mechanizmust a fragmoplaszt nevű struktúra irányítja, amely a hátramaradt interpoláris mikrotubulusokból épül fel a régi mitózis orsó közepén (Alberts et al., 2009).

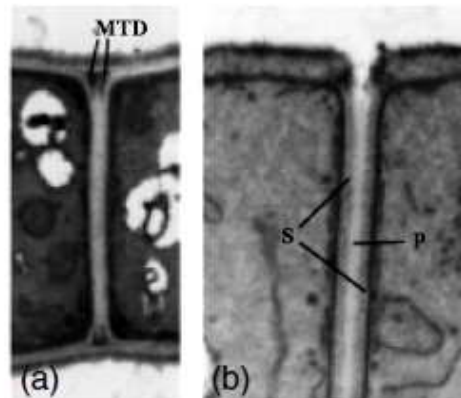
2.5. A hasadó élesztő citokinézise

Ahhoz, hogy a leánysejtek a sejtosztódás után életképesek legyenek, egy háromrétegű (trilamináris) struktúrának kell létrejönnie még a sejtosztódási mechanizmus indulása előtt. A háromrétegű szeptum egy belső, elsődleges szeptumból, és az azt két oldalról körülvevő másodlagos szeptumokból áll (7. ábra).

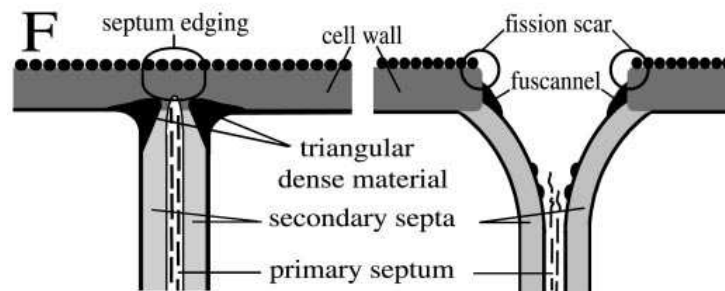
Hasadó élesztő citokinézise során a szeptum két lépésben szintetizálódik. Először a kontraktilis aktomiozin gyűrű jön létre a sejt középvonalában. Az anafázis végén, az aktomiozin gyűrű megjelenését és a leánykromatidák szegregálódását követően egy szeptációt iniciáló hálózatnak (*septation initiation network*, SIN) nevezett jel kaszkád hatására kezd a gyűrű összehúzódni, és ezt kísérve kezd szintetizálódni az elsődleges szeptum. Az elsődleges szeptum két oldalán fokozatosan kezd megjelenni a másodlagos szeptum, a kontraktilis gyűrű mögött (Johnson et al., 1973). A folyamat előrehaladtával a szeptum kettévágja a sejtet, így létre jön a két biológiailag önálló leánysejt. Ahhoz, hogy a két leánysejt fizikailag is elkülönülhessen, az elsődleges szeptumnak és a sejtfal egy keskeny szakaszának le kell bontódnia. Szeptum-hasítási hiánymutások vizsgálata során sok esetben láncszerűen összekapcsolódott sejteket

tapasztaltak, amelyek az el nem bontott szeptumok következtében alakulhattak ki. Ezek a láncszerűen összekapcsolódott sejtek valódi hifák, nem pedig pszeudohifák.

A hifát létrehozó mutánsok genetikai és sejtbiológiai vizsgálatainak segítségével azonosítani tudták a sejtosztódásban résztvevő géneket és fehérjéket. Az elsőként leírt sejtosztódási gén a *sep1* volt, ami egy transzkripciósfaktort kódol.



7. ábra: A szeptum szerkezete *S. pombe* esetében elektronmikroszkópos felvételen (Sipiczki, 2007)



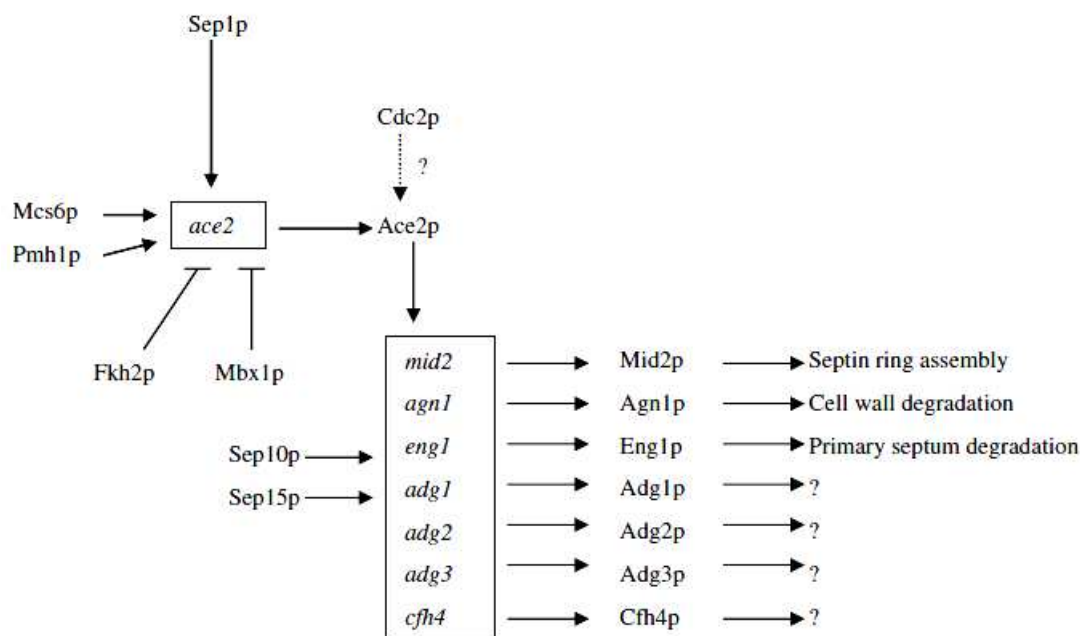
8. ábra: A szeptum vázlatos szerkezete *S. pombe* esetében (Dekker et al., 2004)

Az elsődleges szeptum főleg lineáris 1,3- β -glükánból épül fel. A másodlagos szeptum pedig 1,3- α -glükánból, 1,6- β -glükán elágazásokat tartalmazó 1,3- β -glükánból és 1,6- β -glükánból áll. A sejtosztódást követően a másodlagos szeptum lesz részben a leánysejt sejtfa. Az elsődleges szeptumot, annak a sejtfallal való keskeny érintkezési pontjában, amit szeptum-edging-nek nevezünk, két háromszög-keresztmetszetű struktúra (material triangular dense: MTDs) szegélyezi (8. ábra) (Sipiczki, 2007).

A sejtosztódáshoz szükséges az elsődleges szeptum és a szeptumot körülvevő sejtfalet alkotó anyagok oldódása. Először az elsődleges szeptum élet körülvevő sejtfalet eróziója történik meg az Agn1 jelű 1,3- α -glükánáz által. Az Agn1 egy periodikusan képződő fehérje, amely a szeptumnál található a sejtosztódás közben. Ennek az enzimnek a fokozott kifejeződése a sejtek lízisét okozza, így arra ad következtetést, hogy szükséges a sejtfalet szintézis és annak degradációja közti egyensúly fenntartása a sejtosztódás közben (Martín-Cuadrado et al., 2003). A *S. pombe* rendelkezik egy *agn2* nevű génnel is, amely 1,3- α -glükánázt kódol, de ennek hiánya nem befolyásolja a sejtosztódást. Az Agn2-nek az aszkuszfal endolízisében van szerepe a sporuláció ideje alatt (Dekker et al., 2004).

A sejtfalet bomlását követően a sejt turgor nyomása fizikai erőt fejt ki a másodlagos szeptumra, ezáltal elindul az elsődleges szeptum hasítása. Az Eng1 - amely átmenetileg gyűrű-alakban helyezkedik el a szeptum régióban - segítségével megy végbe az elsődleges szeptum 1,3- β -glükán anyagának bomlása (Sipiczki, 2007). Azonban ennek az enzimnek az aktivitása nem feltétlenül szükséges a sejtosztódáshoz, hiszen az *eng1Δ* hiánymutáns sejtek nem formáltak hifákat, csak az elégtelen sejtosztódás következtében körülbelül négyes csoportokban lazán összekapcsolódva maradnak. Ez az állapot egészen addig maradt fent, amíg a leánysejt el nem kezdett nőni az új végén, és a másodlagos szeptum ki nem emelkedett - hiszen az része lett az új sejtfaletnek -, ekkor ugyanis a lazán kapcsolódott csoport darabjaira esett. Az endoglükánázt kódoló *eng1* a sejtciklus alatt periodikusan fejeződik ki, a sejtben az általa kódolt β -1,3-glükánáz aktivitásának maximuma egybeesik a szeptáció és a sejtosztódás idejével (Martín-Cuadrado et al., 2003). Az *eng2* gén feltételezhetően szintén 1,3- β -glükánázt kódol, azonban ennek hiánya nem okoz sejtosztódási problémát a mutáns sejteknél.

A sejtfalet többi része és a másodlagos szeptum szintén tartalmaz poliszacharidokat, így ezek is a sejtosztódási enzimek célpontjaivá válhatnak, ezért a sejteknek meg kell védeni ezen alkotóit, hogy elkerülje a lízist. A sejtfalet hidrolíziséért felelős enzimek egy többszintes szabályozó rendszer részei (9. ábra), azonban az az elsődleges szignál, amely szükséges az osztódáshoz, nem ismert.

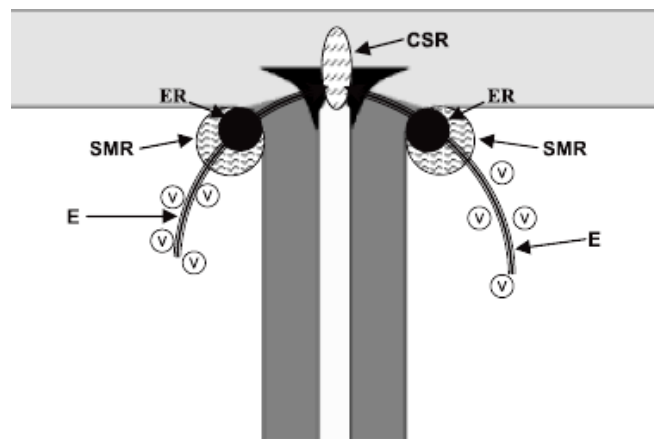


9. ábra: A *S. pombe* sejtosztódását szabályozó rendszer (Sipiczki, 2007)

A *sep1*, az elsőként azonosított sejtosztódási gén, egy fork-head típusú DNS-kötő domént tartalmazó transzkripció faktort kódol (Ribár et al., 1997) (Gould and Simanis, 1997). A sejtosztódás szabályozásában kulcsszerepet játszó *Sep1* maga valószínűleg egy poszt-transzkripció mechanizmus útján szabályozott, mivel a *sep1* transzkripció nem mutat periodicitást a sejtosztódási ciklus alatt. Az első hullám magába foglalja a cink-ujj transzkripció faktort kódoló *ace2* megjelenését, amely felelős a transzkripció második hullámáért. Az *ace2Δ* hiánymutáns sejtek, a *sep1Δ* hiánymutáns sejtekkel megegyező módon hifákat képeznek, ami arra ad következtetést, hogy a *Sep1* az *ace2*-n keresztül kontrollálja a sejtosztódást. A poszttranszlációs foszforiláció által módosított *Ace2*, a *S. cerevisiae*-ben található homológjával ellentétben, szimmetrikusan helyezkedik el a leánysejtek magjaiban (Sipiczki, 2007).

Normál *Eng1* fehérjével rendelkező *agn1Δ* hiánymutáns sejtek esetében az elsődleges szeptum nem bomlott le, ennek lehetséges magyarázata lehet, hogy az *Eng1* nem aktív addig, amíg az elsődleges szeptumot körülvevő sejtfa sértetlen. Lehetséges, hogy a sejtfa lának a fizikai megszakadása által jön létre az a jel, amely az *Eng1* molekulákat aktivizálja.

Ahhoz, hogy az Agn1 és az Eng1 a szeptumban és a szomszédos sejtfa régióban megfelelően helyezkedjen el szükség van szeptinekre, Mid2-re és funkcionális exocitózisra. A Mid2p a citokinézis során képződik a *mid2* gén kifejeződése által, amely a Sep1p-Ace2p szabályozó kaszkád része. Először az összehúzódnáira nem képes szeptin-Mid2-gyűrű (septin-Mid2-ring: SMR) alakul ki a szeptáció oldalán, amely a szeptum megjelenésekor két gyűrűvé hasad. Az exocitózis során a hidrolizáló enzimek a szeptum és a sejtfa elágazási helyein egy sejtosztódási gyűrűt képeznek (cell-separation ring: CSR) (10. ábra), ezen enzimek szállításában és kifejeződésük irányításában vesznek részt az SMR-ek (Sipiczki, 2007).



10. ábra: A sejtosztódásban részt vevő fő komponensek és azok elhelyezkedése a sejtosztódás kezdetén (CSR: cell-separation ring, sejtosztódási gyűrű; E: exocyst system, exocitózis rendszer; v: vezikula, vezikulum; ER: exocyst ring, exocitózis gyűrű, SMR: septin-Mid2-ring) (Sipiczki, 2007)

Mind az Agn1 és Eng1 tartalmaz a szekrécióhoz szükséges szignál szekvenciát. Ahogyan az a 10. ábrán is látható, a CSR, ami az Eng1-t és Agn1-t tartalmazza, a két SMR között helyezkedik el (Sipiczki, 2007).

Ahogy már korábban említettem, a hasadó élesztő sejtjei haploid formában stabilak; aszexuális módon szaporodnak egészen addig, míg a környezetében valamilyen tápanyag hiány fel nem lép. Amikor azonban a sejtek éhezni kezdenek, főleg nitrogén hiány esetében, szexuális módon történő szaporodásra térnek át és ettől kezdve a sejtek az ellentétes párosodási típussal zigótákat alakítanak ki. Ivaros szaporodásuk a sejtek konjugációjával, majd askospórák képzésével történik. Élesztők esetében tehát a meiózist spórák kialakulása követi. Amikor a *S. pombe* diploid sejtjei

meiózison mennek keresztül, akkor négy haploid askospórává válnak szét, az így létrejövő askospórák a környezeti hatásoknak nagyon ellenállóak. Az askospóra falának kialakulásához számos enzim aktivitása szükséges, úgy, mint α - és β -glükán szintetáz. Az askuszfal a vegetatív sejtak falához hasonlóan, α -1,3-glükánt és β -1,3-glükánt tartalmaz. Amikor a spórák már érettek, az askuszfal egy endolitikus folyamaton megy keresztül, amely az askospórák kiszabadulását segíti elő az askuszból, ezáltal a spórák szét tudnak szóródni a környezetben. Ehhez szükséges az endo- α -1,3-glükánázt kódoló Agn2 aktivitása, amely hidrolizálni képes az α -1,3-glükánázt tartalmazó askuszfalat (Dekker et al., 2004). Emellett az Eng2 által kódolt endo- β -1,3-glükánáz jelenléte is szükséges az askospórák kiszabadulásához. Az Eng2 megtalálható a sejtben a vegetatív növekedés során, majd a sporuláció ideje alatt nagy mennyiségben szintetizálódik, melynek maximumát a meiózis II szakaszának végén és a spórák kialakulásakor éri el (del Dedo et al., 2009).

2.6. Filogenetikai analízis és törzsfa szerkesztés

A szekvenciák közti kapcsolatok fontos vizsgálati módja a számítógépes módszerekkel végzett filogenetikai analízis. Segítségével származási kapcsolatok, evolúciókutatások, lehetséges szerkezeti és funkcionális változások követhetők nyomon. Az ilyen módon végzett kutatások azonban valószínűleg nem reprodukálják tökéletesen az evolúciós fákat, amelyek az egyes fajok, gének vagy fehérjék közti evolúciós kapcsolatokat mutatják. Ha közeli fajok vagy nagyon kis gének törzsfáját keressük, akkor célszerűbb nukleotid szekvenciákat vizsgálnunk, mert ezekben minden mutáció látszódik, míg a fehérje szekvenciákat távolabbi fajok evolúciójának kutatására szokták használni. A fehérje általában a génnek egy funkcionális terméke, amelyben lényegesen lassabban jön létre változás, mint a DNS-ben, hiszen az adott fehérje funkciójának megőrzése a szekvencia konzerválódásának hajtóereje. A fehérje alapú filogenetikai analízisek egyik előnye a nukleotidokkal végzett vizsgálatokkal szemben, hogy esetükben nem kell az intronokkal foglalkozni. A különböző DNS régiók – kódoló (exon) és nem-kódoló (intron) szakaszok – ugyanis eltérő sebességgel fejlődnek. Nukleotid szekvenciák analízisét tovább nehezítheti a horizontális géntranszfer előfordulása, illetve a mutáció lehetősége.

A fehérjék hasonlóságának leírására több különböző fogalom ismert. Amennyiben olyan gének csoportjáról van szó, amelyek ugyanazon fajban, egyetlen közös ősi génből alakultak ki, akkor az egyes formákat paralógoknak, vagy egymással paralóg géneknek nevezzük. Ha viszont két különböző faj esetében azonosítható egy olyan gén, amely a két faj egyik közös ősének egyetlen génjéből vezethető le, akkor az említett formákat ortológoknak nevezzük, vagy egymással ortológ géneknek. A homológiának szintén kétféle értelmezése ismert, eredetileg akkor tekintettek két szekvenciát egymással homológoknak, ha a közöttük lévő szekvenciális hasonlóság hátterében az egyazon közös ősi szekvenciából való evolúciós leszármazás húzódtott meg. Újabban viszont a homológiát egyszerűn két szekvencia hasonlóságának leírására alkalmazzák és az evolúciós kapcsolatok kifejezésére inkább a paralógia és ortológia használatát részesítik előnyben (Campbell, 2004). Tehát paralógia esetében gén duplikáció történt a szétválás előtt, míg ortológia esetében a szétválás óta divergencia történt. Minden filogenetikai analízis a lehetséges gén ortológok összehasonlításán alapul.

Ugyanazon genetikai információ-tömegből kiindulva több különböző törzsfá is felállítható attól függően, hogy milyen stratégiával próbáljuk modellezni az evolúciós folyamatokat a rokonsági viszonyok feltárásához. Három különböző megközelítésmódot különíthetünk el: a távolságalapú, a legnagyobb takarékosagra és a legnagyobb valószínűségre törekvő módszereket. A távolságalapú stratégia (Neighbor-Joining) a vizsgált szekvenciák valamennyi elképzelhető párját összehasonlítja, felméri azok távolságát, majd ez alapján hierarchikusan csoportosítva őket rajzolja meg a törzsfát (Saitou and Masatoshi, 1987). A legnagyobb takarékosagra törekvő módszer (maximum parsimony) célja az evolúciós törzsfá ágainak összhosszúságának minimalizálása. Olyan fát keres, amely a lehető legkevesebb evolúciós lépéssel (karakterváltozással, ill. feltételezett mutációval) magyarázza a leszármazási viszonyokat (Jin et al., 2007). A legnagyobb valószínűségre törekvő módszer, vagy más néven maximum likelihood pedig azt a fát keresi, amelyet vizsgálva legnagyobb annak a valószínűsége, hogy az adott szekvenciákból indultunk ki. Sok fát végignévezve kiválasztja a legnagyobb valószínűséghez tartozót (Felsenstein, 1981). Ez a módszer minden pozícióra kiszámítja, hogy adott fa és helyettesítési modell mellett mi a valószínűsége annak, hogy a megfigyelt variációs mintázat jöjjön létre az adott

pozícióban. Az egyes pozíciókra kapott valószínűségek összeszorzásával adódik a teljes fa valószínűsége. Ezt sok fára megnézi, és a legjobbat kiválasztja.

A filogenetikai fák formájában kapott eredmények hitelességének ellenőrzése lehetséges az adatoknak egy másik módszerrel való ismételt kiértékelése, majd a két módszerrel kapott fának az összehasonlításának segítségével. A hitelesség ellenőrzésének egy másik módszere a bootstrap-elemzés, amely során a fa felállítását többször megismétlik oly módon, hogy minden egyes esetben az eredeti szekvencia illesztésből kiindulva az egyes pozíciókat visszatevéses mintavételezéssel megkeverik, és az így kapott mesterséges szekvenciák alapján újra kiszámítjuk a törzsfát, lényegében véletlen mutációkat okozva az adott szekvenciában. Amennyiben a kapott fa szerkezete az adatok ismételt változtatgatásának ellenére is megőrizte alakját, akkor ennek a szerkezetnek a hihetősége megnő (Campbell, 2004).

3. Célkitűzés

TDK munkám alatt a *Schizosaccharomyces pombe* citokinézisében jelentős szerepet betöltő két gén, az *eng1* és a sporulációban érintett *eng2* gének filogenetikai analízisére vállalkozom.

Célom, hogy 30 fajban azonosítsam a *S. pombe*-ben megtalálható sejtosztódási gének közül az *eng1* és az *eng2* lehetséges (feltételezett) ortológjainak előfordulását. A vizsgálandó szekvenciák kiválasztásánál fontos szempont, hogy az adott fajnak a teljes genomi szekvenciája rendelkezésre álljon. Vizsgálataim célpontjában a *Schizosaccharomyces* nemzetség áll, és így az aszkuszos vagy más néven tömlősgombákra fordítom a fő hangsúlyt. A fajok kiválasztásánál azonban külső csoport analíziseként kitekintést teszek a járomspórás és a bazídiumos gombák tagozataiba tartozó fajok közé is.

Az *eng1* gén fontos szerepet tölt be a hasadó élesztő életciklusában, hiszen az általa kódolt fehérje felelős a sejtosztódás során az elsődleges szeptum hasításáért. Az *eng2*-nek azonban nincs szerepe a citokinézisben, viszont a két gén által kódol fehérjék szekvenciáiban lévő nagyfokú hasonlóság miatt lényegesnek tartom a vizsgálatát. Az egyes fajokban mind az *eng1*, mind pedig az *eng2* gének ortológjait fajok közti, majd reciprok páros BLASTp analízissel határozom meg NCBI és Broad adatbázisok használatával.

A feltételezett ortológokat a pontos azonosítás céljából különböző fehérje szerkezetet vizsgáló módszerekkel fogom analizálni. Pfam adatbázis segítségével a szekvenciában fellelhető domének keresését, illetve azonosítását fogom elvégezni. Vizsgálni fogom az *eng1* gén által kódolt fehérjére jellemző szignál peptid jelenlétét az N-terminálison, mind a vélt *eng1* és az *eng2* ortológok esetében. Majd a szekvenciák többszörös illesztését végzem el, ami az evolúciós leszármazások vizsgálatában segít azáltal, hogy a szekvenciák egy csoportja közötti hasonlóságokra, konzerválódott régiókra, biológiai jelentőségű mintázatokra hívja fel a figyelmet. Majd az azonosított fehérje szekvenciák felhasználásával a fajok, pontosabban a két gén (*eng1* és *eng2*) filogenetikai törzsfáját készítem el.

4. Módszerek

4.1. Genomikai adatgyűjtés és adatfeldolgozás

A genomika magába foglalja a DNS megszekvenálását, az egyes populációkon belüli genomi sokféleség, a genomi variabilitás vizsgálatát, és a gének átíródását, transzkripcióját szabályozó rendszerek kutatását. A DNS-szekvenálás során a DNS-t alkotó nukleotidok bázissorrendjét határozzák meg. Minden szervezetben a DNS szekvencia határozza meg az örökletes genetikai információt. Több szekvencia adatbázis is ismert, melyek közül az egyik legjelentősebb a National Center for Biotechnology Information (röviden: NCBI). Az NCBI több, mint 800 szervezet teljes genomi szekvenciáját és számos, részben megszekvenált szervezet információit tartalmazza, ezzel a bioinformatikai kutatások alapját képezve. Egy másik bioinformatikai adatbázist szolgáltat a Broad Institute. Mindkét adatbázis online, bárki számára hozzáférhető és az adatokon alkalmazható funkciók széles körével rendelkezik.

Munkám során 30 kiválasztott faj (1. táblázat) fehérje szekvenciáján végeztem vizsgálatokat, a *Schizosaccharomyces* csoport kivételével minden szekvenciát az NCBI adatbázisból nyertem ki. Az NCBI adatbázis a vizsgálataim központjában álló *Schizosaccharomyces* csoportba tartozó négy élesztő közül csak kettőnek: a *S. pombe*-nek, és a *S. japonicus*-nak tartalmazta a szekvenciáját, a Broad Institute által kínált adatbázissal szemben, amely a további két hasadó élesztőnek (*S. octosporus* és *S. cryophilus*) is tartalmazta az általam keresett szekvenciáját, ezért végül mind a négy élesztő faj szekvenciáját a Broad Institute adatbázisából használtam.

1. táblázat: A vizsgált fajok listája

Fajok	Törzs	Adatbázis
<i>Ashbya gossypii</i>	ATCC 10895	NCBI
<i>Aspergillus flavus</i>	NRRL3357	NCBI
<i>Aspergillus fumigatus</i>	Af293	NCBI
<i>Aspergillus nidulans</i>	FGSC A4	NCBI
<i>Blastomyces dermatitidis</i> (<i>Ajellomyces dermatitidis</i>)	ATCC 18188	NCBI
<i>Botrytis cinerea</i> (<i>Botryotinia fuckeliana</i>)	B05.10	NCBI

Az 1. táblázat a 27. oldalon folytatódik

1. táblázat: A vizsgált fajok listája (folytatása)

Fajok	Törzs	Adatbázis
<i>Candida albicans</i>	WO-1, SC5314	NCBI
<i>Candida glabrata</i>	CBS 138	NCBI
<i>Candida tropicalis</i>	MYA-3404	NCBI
<i>Coccidioides immitis</i>	RS	NCBI
<i>Cryptococcus neoformans</i>	JEC21	NCBI
<i>Debaryomyces hansenii</i>	CBS767	NCBI
<i>Histoplasma capsulatum</i> (<i>Ajellomyces capsulatus</i>)	G186AR, H88, NAm1, H143	NCBI
<i>Kluyveromyces lactis</i>	NRRL Y-1140	NCBI
<i>Magnaporthe oryzae</i>	70-15	NCBI
<i>Neosartorya fischeri</i>	NRRL 181	NCBI
<i>Neurospora crassa</i>	OR74A	NCBI
<i>Paracoccidioides brasiliensis</i>	Pb01, Pb03, Pb18	NCBI
<i>Penicillium chrysogenum</i>	Wisconsin 54-1255	NCBI
<i>Phycomyces blakesleeanus</i>	NRRL1555	NCBI
<i>Rhizopus oryzae</i>	RA 99-880	NCBI
<i>Saccharomyces cerevisiae</i>	EC1118, RM11-1a	NCBI
<i>Schizosaccharomyces cryophilus</i>	NRRL Y-48691T	Broad Institute
<i>Schizosaccharomyces japonicus</i>	yFS275	Broad Institute
<i>Schizosaccharomyces octosporus</i>	yFS28	Broad Institute
<i>Schizosaccharomyces pombe</i>	ATCC 38366/972h ⁻	Broad Institute
<i>Trichoderma reesei</i>	QM6a	NCBI
<i>Ustilago maydis</i>	FB1	NCBI
<i>Verticillium albo-atrum</i>	VaMs.102	NCBI
<i>Yarrowia lipolytica</i>	CLIB122	NCBI

*A táblázatban szereplők adatokat 2011.09.17-én töltöttem le

A vizsgálandó szekvenciák kiválasztásánál fontos szempont volt, hogy az adott fajnak a teljes genomi szekvenciája rendelkezésre álljon. Vizsgálataim során az aszkuszos vagy más néven tömlősgombákra fordítom a fő hangsúlyt, de a fajok kiválasztásánál kitekintést teszek a járomspórás (*Phycomyces blakesleeanus*, *Rhizopus oryzae*) és a bazídiumos gombák (*Cryptococcus neoformans*, *Ustilago maydis*)

tagozataiba tartozó fajok közé is. A fajok kiválasztásának alapját H. Wang 2009-ben megjelent cikke adta, amelyben több, mint 80 gombfaj filogenetikai törzsfájának elkészítését írja le (Wang et al., 2009).

A proteomika a proteom, vagyis az élő szervezetben előforduló összes fehérjének a megismerésével foglalkozik. Aminosav szekvenciának a fehérjék elsődleges szerkezetét nevezzük. A fehérjék aminosav sorrendje a gének nukleotid szekvenciájában kódolt. A proteomika témakörébe tartozik a fehérjék szerkezetének, biológiai funkciójának és ezek térbeli és időbeli változásának megismerése. Vizsgálódási körébe tartozik továbbá a fehérje eredetének meghatározása, rendszertani besorolása, a különböző forrásból származó, de azonos biológiai funkciót ellátó proteinek szerkezetének összehasonlítása, a fehérje jelenlétének vagy hiányának igazolása.

Mind az NCBI, mind pedig a Broad Institute által szolgáltatott internetes adatbázisokon megtalálható a BLAST (Basic Local Alignment Search Tool) nevű algoritmus, mellyel végzett keresés eredményeként olyan találatokat kapunk, amelyek az adott adatbázisban található génszekvenciák közül szignifikáns hasonlóságot mutatnak az általunk keresett szekvenciával. A BLAST heurisztikus módszert alkalmaz, melynek lényege az, hogy az adatbázisból kiválasztja azokat a szekvenciákat, amelyek valószínűleg hasonlóak a célszekvenciánkhoz és csak ezeknek a hasonlóságát vizsgálja. Heurisztika alkalmazására azért van szükség, mert nagyméretű adatbázis esetében annak minden egyes eleméhez való hasonlítás nagyon sok ideig tartana. Megkülönböztethetünk nukleotid (BLASTn) és fehérje (BLASTp) szekvenciák összehasonlításával foglalkozó algoritmusokat. Munkám során a *S. pombe*-ben található *eng1* és *eng2* gének által kódolt fehérjék szekvenciájából kiindulva végeztem BLASTp típusú vizsgálatokat.

Egy ilyen találat hasonlóságának szignifikancia szintjét jellemzi az E-érték (expect value). Minden kapott találatnak vannak a keresett szekvenciával egy az egyben megfeleltethető szakaszai, ezek alapján számítható ki az úgy nevezett bit score. Az E-érték azt mutatja meg, hogy az adott keresésnél statisztikailag várhatóan hány, a megadott bit score-értékkel megegyező vagy annál nagyobb bit score-ú hasonlóságot mutató egyezés lesz a véletlen műve, és nem valamilyen evolúciós kapcsolat jele. Általánosságban kimondható, hogy minél nagyobb az E-értéke, annál valószínűbb, hogy a hasonlóság csak a véletlen műve. Az E-értékét három tényező szabja meg: a bit score,

a keresett szekvencia hossza és az adatbázis mérete. A nagyobb bit score értékek csökkentik az E-értékét, hiszen minél nagyobb a hasonlóság két szekvencia között, annál kisebb a valószínűsége, hogy az csak a véletlen műve lenne. A hosszabb szekvenciákra való keresés, azonban növeli az E-értéket, mivel ugyanaz a bit score érték, hosszabb vizsgált szekvencia esetén könnyebben lehet a véletlen műve. Azonban minél nagyobb az adatbázis, annál könnyebb benne olyan szekvenciát találni, amely hasonlít az általunk keresett szekvenciára, ezért az adatbázis méretének növelése, növeli az E-értéket. Az E maga az $E=mn2^{-S}$ képletből adódik, ahol m a keresett szekvencia effektív hossza, n az adatbázis effektív hossza, bázisokban kifejezve, S pedig a bit score érték (Campbell, 2004).

Mindkét adatbázis használata lehetőséget ad egyéni beállítások alkalmazására. Az NCBI által szolgáltatott BLASTp esetében állítani lehet: a maximális számú igazított sorozatok megjelenítését (Max target sequences), vagyis hogy mekkora szekvenciákat vegyen észre az algoritmus, továbbá a Word size állításával szabályozható, hogy az illesztés során mekkora egységekre darabolja a szekvenciát. Megválasztható az E küszöbértéke, az alkalmazott szubsztitúciós mátrix, a gap-ekért adott pontszám, és állítani lehet még a kompozíciós kiigazítást, illetve filterek és maszkok használatát. Filter használatával beszűkíthetjük a találatokat azokra, amelyek nagyon hasonlóak, ennek például domén kereséskor lehet jelentősége. Broad Institute által szolgáltatott programmal végzett BLASTp vizsgálatok esetében szintén lehetőség van bizonyos beállítások módosítására, azonban az itt állítható paraméterek száma lényegesen kevesebb, mint az NCBI adatbázis esetében. Munkám során a BLASTp típusú analíziseket az NCBI adatbázison végeztem, többek között az általa kínált nagyobb paraméterállítási lehetőségek miatt, illetve azért, mert a Broad Institute adatbázison kapott BLASTp eredmények az algoritmus által a szekvenciába rakott gap-ek számát nem mutatta meg, amit azonban fontos adatnak tekintettem. Továbbá a Broad Institute BLASTp eredményében az E-értékek néhány esetben 0-nak adódtak, ami ellentmondott az NCBI által azonos keresési feltételek mellett kapott E-érték jelentésének. Azt feltételezem, hogy mind a Broad Institute által kínált, mind pedig az NCBI-on elérhető BLASTp esetében van egy kritikus E-érték, amely felett a program 0-t ír ki. A két adatbázis esetében a küszöbérték azonban eltérő, így ez megtévesztő lehet. Ezért vizsgálataimat az NCBI által szolgáltatott programmal végeztem és az általa felkínált alapbeállításokat alkalmaztam, tehát a maximális számú igazított sorozatok

megjelenítésének száma 100 volt, az E küszöbértéke 10, BLOSUM62 mátrixot használtam. BLOSUM mátrixban a szekvenciákból csoportokat, klasztereket képeznek azok hasonlósága alapján, például egymással 62%-nál nagyobb azonosságot mutató szekvenciák egy klaszterbe kerülnek. Továbbá egy új gap megjelenéséért 11 pont, egy már meglévő gap hosszabbodásáért pedig 1 pont járt. Feltételes kompozíciós pont mátrix kiigazítást hagytam kiválasztva a lehetőségek közül, filtereket és maszkokat pedig nem használtam.

4.2. Fehérje szerkezet vizsgálat

Irodalmi adatokból ismert, hogy az általam vizsgált *eng1* gén által kódolt fehérje szekvenciája tartalmaz a szekrécióhoz szükséges szignálpeptidet, ellentétben a másik általam vizsgált, *eng2* génnel szemben (del Dedo et al., 2009). A szignálpeptid átlagosan 20 aminosavas, erősen hidrofób szekvencia a fehérje elsőként szintetizálódó N-terminális végén, amelynek feladata a szintézist végző riboszomális komplexeknek a citoplazmából a durva felszínű endoplazmás retikulumra való irányítása (Campbell, 2004).

Több számítógépes program áll rendelkezésünkre, amelyek lehetőséget adnak a fehérjék azonosítására. Ezek közül az egyik, a széles körben alkalmazott Pfam (Protein families database of alignments and HMMs) adatbázis, amely konzervált fehérje családokat és doméneket tartalmaz. Használatával lehetséges új szekvenciák esetében a fehérje funkciójának megállapítása, illetve a fehérje család meghatározása. A Pfam egy rejtett-Markov modelleket (hidden Markov model: HMM) alkalmazó adatbázis. Két változata ismert: a Pfam-A, amely kurátorok által annotált doméneket tartalmaz, és a Pfam-B, amely automatikusan generált doméneket tartalmaz. Az adatbázis használata során alapbeállításként a Pfam-A-ban lehetséges a keresés, de egyéni beállítások során a Pfam-B-re is kiterjeszthető a munka (Finn et al., 2010).

Egy másik adatbázis, amely specifikusan a szekretált fehérjék azonosítására használható a SignalP 3.0. A program a kérdéses aminosav szekvencia elemzésére két módszert használ, hogy a szignálpeptidek jelenlétét megjósolja: a Neural network (röviden: NN) és a rejtett-Markov modellt. A program használatakor lehetőségünk van a beállítások változtatására, többek között meg kell adni, hogy a kérdéses szekvencia az eukarióta, Gram-pozitív vagy Gram-negatív csoportba tartozik-e, vizsgálataim során az

eukarióta szervezeteket állítottam be. Továbbá választhatunk, hogy a keresett szekvencia tartalmazhat-e transzmembrán régiót vagy sem, esetemben előfordulhat, hogy az aminosav szekvencia tartalmaz ilyen régiót. Beállíthatjuk, hogy az eredményeket milyen grafikus képi formátumban kívánjuk látni, illetve, hogy az eredményül kapott információkat milyen részletességig szeretnénk megkapni. A program használata során az alapbeállításokat alkalmaztam. A NN által kapott grafikus eredmény három különböző pontszámot is megad: a C-t, az S-t és az Y-t. Két további pontszám is látható: az S-középérték és a D pontszám. Két különböző NN-t használ a program, az egyik megjósolja a szignálpeptidet, a másik pedig a szignál peptidáz I. hasítási helyét próbálja megjósolni. Az S-érték a szekvenciában szereplő minden egyes aminosavra ki van számolva, az egyes aminosavak esetében a magas pontszám jelzi, hogy a kérdéses aminosav része a szignálpeptidnek; alacsony pontszám pedig azt jelenti, hogy az érett fehérje része az adott aminosav. A C-érték (cleavage site-score) vagy hasadási hely-érték, a benyújtott szekvencia minden egyes részére kiszámított, de csak a tényleges hasadási helyen magas a pontszáma. Megadja, hogy a szekvencia hányadik aminosava lesz az érett fehérje első eleme. Az Y-max érték a C- és az S-értékekből származtatott mennyiség, amely egy biztosabb előrejelzést ad a hasadási helyet illetően, mint a C-érték önmagában. Az S-középérték az S-pontszámok átlagát mutatja és a kérdéses szignál peptid hosszát hívatott jelezni. A D-érték pedig az S-középértékek átlagából és az Y-értékből számított mennyiség. Ha a kérdéses fehérje szekvencia nem tartalmaz szekrécióhoz szükséges szignálpeptid részt, akkor az előbb felsorolt mennyiségek mindegyike alacsony értéken marad (Bendtsen et al., 2004). A rejtett-Markov modell használata pedig lehetővé teszi a valódi szignál peptidek és a behorgonyozott szignálok megkülönböztetését. A nem valódi szignál peptidek, olyan fehérje szekvenciák, amelyek szintén az adott fehérje transzlokációját segítik elő, de ezek nem kerülnek hasításra a szignál-peptidáz által. Ahogyan a polipeptid lánc a membránon keresztül transzlokálódik, a fehérje hidrofób része a membránhoz horgonyozva marad (Nielsen et al., 1999). Az NN és HMM módszerek együttes alkalmazása jelentős részben kezelni tudja a szekvenciák sokféleségét.

4.3. Többszörös szekvencia illesztés

A többszörös szekvencia illesztés lehetővé teszi egy-egy géncsaládon vagy fehérjecsaládon belül az összefüggések felismerését és a családra jellemző, konzerválódott, biológiailag jelentőséggel bíró mintázatok feltárását. Továbbá segít az új és meglévő szekvenciák közötti hasonlóságok felderítésében, a fehérjék másodlagos és harmadlagos szerkezetének előrejelzésében, illetve filogenetikai analízisekhez szükséges. A páronkénti összerendezéshez képest növeli a jel/zaj viszonyt és segítségével olyan finom hasonlóságokat vehetünk észre, amiket a páronkénti illesztés nem mutat ki.

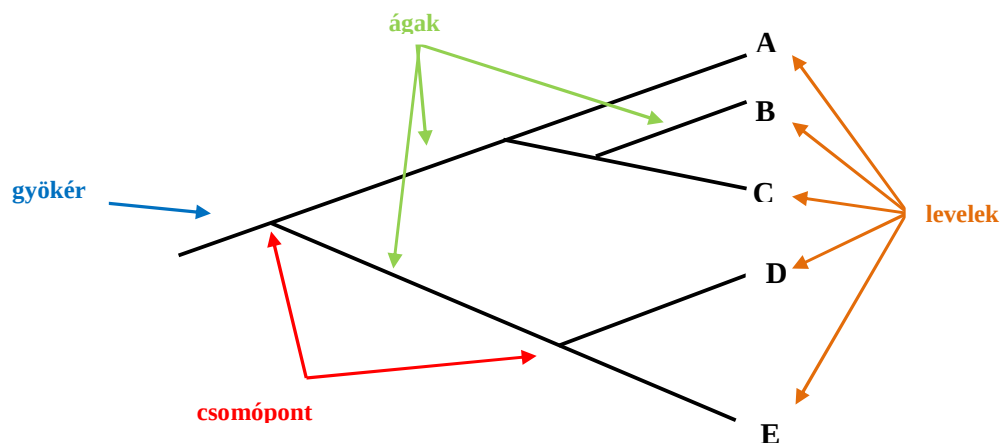
A legtöbb program a progresszív illesztés módszerét alkalmazza, amelynek egyik fontos előnye, hogy akkor is megfelelő eredményt ad, ha eltérő a szekvenciák hasonlósága, emellett a módszer gyors és érzékeny. Az általam használt Clustal nevű program is a progresszív illesztés módszerét alkalmazza a többszörös illesztés elvégzéséhez. A ClustalW nevében a 'W' az angol 'weighted', vagyis súlyozott szót jelenti, a program ugyanis az illesztés különböző részeit különbözőképpen súlyozza. Az összerendezés pontszámának kiszámításakor a közeli szekvenciákat kisebb súllyal veszi figyelembe, hogy a közel azonos szekvenciák ne nyomják el a távolabbiakat. A progresszív illesztés folyamata három lépésre tagolódik, első lépésként páronkénti illesztések konstruálása és távolságmátrix létrehozása történik meg. Az összes lehetséges pár illesztésével hozzuk létre a hasonlósági mátrixot, amelyben a hasonlóságot a pontos megegyezés és a szekvencia hosszának hányadosa adja (azonossági százalék). Ezt követően egy Neighbor-joining irányító fa felépítése következik, majd pedig egy progresszív illesztés történik az irányító fa vezérlésével. A progresszív illesztést a két leghasonlóbb szekvenciával kezdjük, utána az irányító fát követve egyenként illesztjük a szekvenciákat a már meglévő eddigi illesztéshez, a program szükség szerint hézagokat vezet be a fehérje aminosav sorrendjébe. A ClustalW grafikus felülete a ClustalX nevű program (Mount, 2004).

4.4. Törzsfa készítés

A filogenetikai analízis célja az evolúciós viszonyok, az evolúciós történet feltárása, az ilyen jellegű kutatások eredménye általában egy filogenetikai fa (törzsfa).

A filogenetikai törzsfákat építő módszerek egy bizonyos evolúciós modell érvényességét feltételezik. Legfontosabb feltételezése az, hogy a törzsfejlődés fával ábrázolható. Ez nem mindig igaz, pl. hibrid fajok, illetve mikroorganizmusok között előforduló DNS-átvitel esetén, továbbá megjelenhet horizontális géntranszfer is, amely előfordulása szintén nehezíti az ilyen irányú kutatásokat.

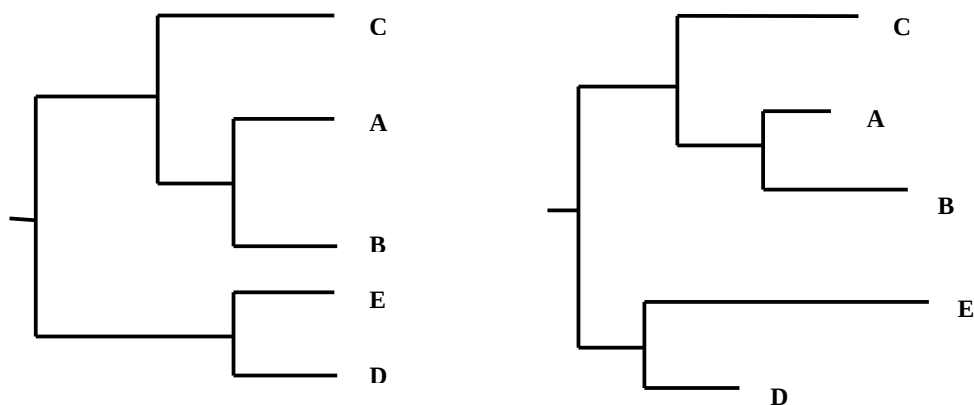
Egy törzsfa vázlatos rajza látható a 10. ábrán. A levelek vagy OTU-k = operational taxonomic unit, legtöbbször ma élő faj, de lehet kisebb vagy nagyobb taxonómiai egység is. A csomópontok vagy elágazódások (nóduszok), ősi fajokat reprezentálnak, szokás HTU-nak (hypothetical taxonomic unit) is nevezni őket.



10. ábra: A törzsfa részei

Megkülönböztetünk gyökeres és gyökér nélküli törzsfákat. A 10. ábrán egy gyökérrel rendelkező filogenetikai fa látható, amelyen a gyökértől minden levélhez vezet egy evolúciós irányt mutató útvonal. Gyökér nélküli filogenetikai fák esetében nem ismert az evolúció iránya, az adott törzsfa csak a rokonsági kapcsolatokat jeleníti meg. Ha az egyik belső pontot kitüntetjük, mint legősibb (múltbéli) fajt, akkor ún. gyökereztetett (rooted) fáról beszélünk, és a kitüntetett pontot hívjuk a fa gyökerének. Ha nincs ilyen pont, akkor a fa gyökerezetlen (unrooted) (Mount, 2004). Több filogenetikai fa-ábrázolási mód is használatos (11. ábra).

Ultrametrikus törzsfa esetében a csomópontok helyzete a szétválások idejét mutatja, ilyenkor az egyes fajok egy oszlopban helyezkednek el, ebben az esetben az ághosszak nem tükrözik az evolúciós távolságot. Ellentétben az additív törzsfákkal, melyeknél az ághosszak az evolúciós távolságokat fejezik ki. A pontok kapcsolódásának módját a fa topológiájának hívjuk.



11. ábra: Ultrametrikus és additív fa ábrázolások

A törzsfa készítésre alkalmas algoritmusok nukleotid szekvenciákat vagy a fehérjék aminosav szekvenciáit használják. A legismertebb faépítő módszerek többszörös szekvencia illesztést alkalmaznak. A PHYLIP (Phylogeny Inference Package) nevű program a filogenetikai analízisek széleskörű módszereinek gyűjteményével rendelkezik. A csomag tartalmazza a legnagyobb takarékoságra, a legnagyobb valószínűsége és a távolságalapú filogenetikai fák előállításának lehetőségét is, valamint bootstrap és konszenzus fa létrehozására is alkalmas (Felsenstein, 1993).

A PHYLIP programban a bootstrap-elést követően az adatok kinyerhetők és egy másik törzsfa rajzoló programba átvihetők. A bootstrap-pelt adatokat a Treewiew nevű program segítségével ábrázoltam. Munkám során távolságalapú (Neighbor Joining) törzsfát készítettem, majd az így létrehozott törzsfa jóságát bootstrap módszerrel ellenőriztem, végül a kapott filogenetikai fát ábrázoltam.

5. Eredmények

5.1. Szekvencia illesztés

Munkám kezdetén, annak érdekében, hogy a két kiválasztott sejtosztódási gén által kódolt fehérjét jobban megismerjem, adatokat, információkat gyűjtöttem az Eng1 és Eng2-ről. Kíváncsi voltam többek között arra, hogy az adott fehérje pontosan milyen funkciót tölt be a sejt életében, van-e katalitikus aktivitása, illetve poszttranszlációs módosulása, milyen hosszú a fehérje szekvencia, és hogy hogyan fejeződik ki a fehérje a sejtciklus alatt. Az adatok gyűjtésében az Uniprot és a Sanger Institute gondozásában álló GeneDB adatbázisok voltak segítségemre.

A fehérjék hasonlóságának leírására több különböző fogalom ismert. Paralógnak nevezzük olyan gének csoportját, amelyek ugyanazon fajban megtalálhatóak, és egyetlen közös ősi génből alakultak ki. A *S. pombe*-ben jelenlevő *eng1* és *eng2* gének egymás paralógjai. Ortológoknak nevezzük két különböző faj esetében azonosítható olyan gének csoportját, amelyek a két faj egyik közös ősének egyetlen génjéből vezethetők le. Paralógia esetében és így az általam vizsgált sejtosztódási gének esetében is gén duplikáció történhetett a szétválás előtt, míg ortológia esetében a szétválás óta történt divergencia.

A szekvencia illesztés meghatározza az illesztett szekvenciák közötti azonosságot, illetve eltérés mértékét. Ortológ szekvenciák esetén a különböző fajokból származó szekvenciák közötti eltérések a fajok független evolúciós múltja során gyűjtött mutációk eredményei. A mutációk száma, vagyis a szekvencia eltérés foka közelítőleg megfeleltethető a két faj közti evolúciós távolsággal: minél régebben vált el egymástól két faj, ortológ génjeik között annál nagyobb lesz a szekvencia eltérés.

A *S. pombe eng1*, majd pedig *eng2* génjéből kiindulva először fajok közötti BLASTp analíziseket végeztem a 30 kiválasztott fajban. Kereső szekvenciának (Query Sequence) mindig a *S. pombe eng1*, illetve *eng2* génje által kódolt fehérje szekvenciáját írtam be az NCBI adatbázisba, a keresési beállításoknál pedig az éppen vizsgált faj nevét választottam ki, ezzel is szűkítve az adatbázis méretét. A szűkítésre azért volt szükség, mert a program az egész adatbázisban való keresés közben nagyon sokáig futott, illetve néhány esetben nem jelenített meg minden találatot. Valószínűleg van egy küszöbérték, ami utáni találatokat nem jelenít meg egy oldalon. Az így kapott eredmények közül, minden egyes fajban megtalált lehetséges ortológok közül

kiválasztottam a legjobban hasonlító szekvenciákat. A megfelelő találatok kiválasztásánál fontos szempont volt az E-érték, hiszen minél kisebbnek adódott az E, annál valószínűbb, hogy a szekvenciák hasonlósága nem a véletlen műve. Emellett figyeltem a talált szekvencia hosszát, illetve hogy azon belül milyen hosszú szakaszra történt az illesztés. A BLASTp eredményében megjelenik az adatbázis által a keresett és a hasonlóknak vélt szekvencia azonosítására szolgáló kód, illetve az adott génnek által kódolt fehérjék leírására vonatkozó adatok, továbbá a fehérje hossza, a pontszámok (bit score és raw score), a találat statisztikai valószínűségét mutató E-érték, a teljes mértékben azonos aminosavak száma és aránya (identities), a pozitív, vagyis hasonló találatok száma és aránya (positives), és adott esetben az illesztés során a szekvenciába helyezett szünetek, gap-ek száma. Emellett megjeleníti az illesztést is, ahol a felső sorban található a kereső szekvencia, alatta pedig a talált szekvencia; a két szekvencia között pedig látható, hogy hol talált az algoritmus azonos (adott aminosav betűkódja) vagy hasonló („+” jelölve) aminosavakat, esetleg hova rakott gap-et („-” jelölve).

Az *eng1* és *eng2*-ből kiinduló BLASTp eredményeket táblázatos formában gyűjtöttem össze. A 2. táblázatban a *S. pombe eng1* génjéből kiinduló BLASTp eredmények egy részlete látható. Az eredményül kapott táblázatokat a Függelékben tüntettem fel. Minden egyes megtartott szekvencia találat esetében a táblázatban feltüntettem a fehérje nevét, amennyiben az NCBI-ban találtam erre vonatkozó adatot, a fehérje hosszát, az azonos és hasonló aminosavak számát, a bit- és raw-score értékeket, és az E-értékét.

Néhány faj esetében egynél több szekvenciát hagytam meg megfelelően jó találatnak, figyelembe véve, hogy az adott szekvencia lehetőleg a teljes genom megszekvenálásából adódjon. Egyes fajokban viszont annyira rossznak ítélt meg a találatokat, hogy a további elemzésekből kihagytam őket, mert a kapott eredmények alapján nem valószínűsíthető az *eng1* és *eng2* ortológok jelenléte. Rossznak ítélt meg a BLASTp találatát, amikor az E-érték kiugróan nagy számnak adódott a többi eredményhez képest, általában az E-érték nagyságrendileg a 10^{-100} vagy ennél kisebb számnak adódott, rossznak ítélt esetben pedig például 10^{-10} nagyságrend is előfordult. A járomspórások tagozatába tartozó *Phycomyces blakesleeanus* és *Rhizopus oryzae*, illetve a bazídiumos gombák tagozatába tartozó *Cryptococcus neoformans* és *Ustilago maydis* fajokat nem vettem alá további analíziseknek. Ezen találatok esetében a BLASTp a keresett génnek szekvenciájához és a többi megtalált lehetséges ortológ szekvenciához

képezt is nagyon rövid szakaszokra illesztette a szekvenciákat, és az E-értékek is kiugróan nagy számoknak adódtak, vagyis nagyon kicsi annak a valószínűsége, hogy a megtalált szekvencia valóban az általam keresett gén ortológja lenne. Az általam vizsgált aszkuszos gombafajok mindegyikében azonosítani tudtam az *eng1* és *eng2* gének ortológjait, egyéb gombafajok esetében azonban nem találtam lehetséges ortológokat.

2. táblázat: *S. pombe eng1* génjének BLASTp eredményei a *Schizosaccharomyces* nemzetségen belül

Faj	<i>S. pombe</i>		<i>S. japonicus</i>		<i>S. octosporus</i>		<i>S. cryophilus</i>	
Fehérje	Eng1	Eng2	Eng1	endo-1,3(4)-beta-glucanase	Eng1	Eng2	Eng1	Eng2
Fehérje hossza (aa)	1016	706	1008	681	1043	707	1055	706
pos		417/699 (60%)	733/1014 (72%)	393/688 (57%)	746/1037 (72%)	415/699 (59%)	750/1057 (71%)	413/710 (58%)
ident		287/699 (41%)	578/1014 (57%)	290/688 (42%)	591/1037 (57%)	286/699 (41%)	593/1057 (56%)	290/710 (41%)
gaps		24/699 (3%)	14/1014 (1%)	34/688 (5%)	48/1037 (5%)	22/699 (3%)	51/1057 (5%)	24/710 (3%)
score		513 bits (1320)	1098 bits (2840)	516 bits (1329)	1189 bits (3077)	528 bits (1359)	1186 bits (3068)	524 bits (1349)
E		2,00E-149	0.0	2,00E-150	0.0	7,00E-154	0.0	1,00E-152
reciprok BLASTp első találat	Fehérje	Eng2	Eng1	Eng2	Eng1	Eng2	Eng1	Eng2
reciprok BLASTp második találat		Eng1	Eng2	Eng1	Eng2	Eng1	Eng2	Eng1

pos jelöli a hasonló aminosavak, ident jelöli a teljesen azonos aminosavak számát

E-értékét normál alakban értelmezve pl. 2,00E-149=2*10⁻¹⁴⁹

Ahogy az a 2. táblázatban is látható, a lehetséges ortológoknak vélt szekvenciákkal reciprok BLASTp analízist végeztem, ami azt jelenti, hogy a talált szekvenciát használtam kereső szekvenciának, ilyenkor arra voltam kíváncsi, hogy a visszakeresésnél az *eng1* vagy *eng2* gének által kódolt fehérjék jönnek-e vissza első találatra. A reciprok BLASTp vizsgálatokat csak a *S. pombe eng1* és *eng2* génjeire kerestem vissza. Amennyiben például az *eng1*-ből kiinduló fajok közötti BLASTp analízist követő visszakeresés eredményeként az *eng1* gént kaptam vissza első találatnak az *S. pombe*-ben található paralógok közül, akkor azt a találatot nagyobb

valószínűséggel nyilvánítottam *eng1* ortológoknak. A vizsgált fajok többségében találtam *eng1* és *eng2* lehetséges ortológokat is, de előfordult olyan eset is, amikor egynél több feltételezhetően *eng2* ortológot találtam egy fajon belül, például a *Yarrowia lipolytica* esetében egy törzsön belül három szekvencia is *eng2* ortológoknak bizonyult. A három szekvenciát vizsgálva az E-értékeket minden esetben elegendően kicsinek találtam, és a kérdéses szekvenciák közti hosszkülönbségek alapján is három különböző génnek tekintem őket, ezért valószínűsíthető, hogy a *Yarrowia lipolytica*-ban az *eng2* gén paralógjai vannak jelen. Több faj vizsgálata során egynél több törzsben is találtam lehetséges ortológokat, *Saccharomyces cerevisiae* esetében például két törzs teljes genom szekvenciájában, *Paracoccidioides brasiliensis* esetében pedig három törzs szekvenciájában is azonosítani tudtam a keresett gének feltételezett ortológjait (Függelék 1 és 2.).

5.2. Fehérje domének azonosítása

Annak érdekében, hogy biztosabb döntést tudjak hozni az egyes ortológokat illetően, a talált szekvenciákat fehérje szerkezeti vizsgálatoknak vettem alá. Először a Pfam adatbázist használtam a lehetséges domének előfordulásának felkutatására, ahol Pfam-A analíziseket végeztem. Ennek eredményeként elmondható, hogy minden általam vizsgált fajban talált lehetséges gén ortológ által kódolt fehérje a glikozil hidroláz 81 fehérje családnak tartozik. A *Schizosaccharomyces* csoporton belüli tagok mindegyikénél az *eng1* ortológok esetében a domén kereső három szénhidrogén kötő domént talált a szekvenciákon belül (12. ábra). Egyéb fehérje domént a Pfam azonban



nem talált.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		Bit score	E-value
				Start	End	Start	End	From	To		
Glyco_hydro_81	Glycosyl hydrolase family 81	Family	n/a	57	733	61	733	5	695	647.6	1.5e-194
Carb_bind	Carbohydrate binding	Domain	n/a	790	841	791	841	2	52	85.9	9e-25
Carb_bind	Carbohydrate binding	Domain	n/a	879	930	879	929	1	51	81.5	2.2e-23
Carb_bind	Carbohydrate binding	Domain	n/a	965	1016	966	1013	2	49	83.2	6.6e-24

12. ábra: Pfam-A fehérje domén kereső adatbázis eredménye *S. pombe eng1* génjét vizsgálva

Az *eng1* gén esetében az irodalmi adatokból ismert szignálpeptidek jelenlétét azonban a Pfam nem tudta igazolni, ezért egy specifikusan szignálpeptid kereső programmal, a SignalP 3.0-val folytattam további analíziseket. A *Schizosaccharomyces* csoporton belül a program egyértelműen igazolni tudta (egy eset kivételével) az irodalmi adatokat, miszerint az *eng1* gén által kódolt fehérje tartalmaz a szekrécióhoz szükséges szignálpeptidet, az *eng2* által kódolt fehérje viszont nem. Fontos megjegyezni azonban, hogy *S. japonicus eng2* szekvenciáját vizsgálva a SignalP 3.0 eredménye nem volt egyértelmű, mert az általa használt kétféle algoritmus eltérő eredményt adott. A rejtett-Markov modell az irodalmi adatokkal összhangban nem talált szignálpeptidet, az NN eredményét tekintve azonban ezzel ellentétben az öt pontszám (C-, Y-, S-, S-középérték és D-pontszám) közül négy esetben (Y-, S-, S-középérték és D-pontszám) szignálpeptid jelenlétére utal. A további fajokban talált ortológokat is vizsgálva azonban több esetben előfordult, hogy a reciprok BLASTp eredményeként feltételezhető *eng2* ortológ rendelkezett szignálpeptiddel (például *Saccharomyces cerevisiae* és *Blastomyces dermatitidis*), illetve olyan esetek is voltak, amikor a feltételezett *eng1* ortológnál nem azonosított szignálpeptidet a program. Az utóbbi eset fordult elő az *Ashbya gossypii* és a *Paracoccidioides brasiliensis* fajoknál. A reciprok BLASTp eredményeit a SignalP 3.0 eredményeivel összevetve meglehetősen színes képet kaptam a két gén pontos azonosítását illetően, ezért jelen dolgozatomban nem vállalkoztam az *eng1* és *eng2* ortológok elkülönítésére a szignálpeptid meglétének vagy hiányának igazolásával. Emellett tovább nehezíti a dolgot a két gén által kódolt fehérje szekvenciája közti nagyfokú hasonlóság megléte. A szekvenciák hosszát illetően a *S. pombe eng1* génje által kódolt fehérje hossza: 1016, *eng2* génje által kódolt fehérje hossza: 706 aminosav hosszú, a vizsgált gombafajokban az ortológ szekvenciák hossza pedig gyakran e két érték közé esik, ami szintén korlátozza a pontos azonosítást. Valószínűsíthető, hogy a két gén ortológját a legtöbb esetben nem lehet egyértelműen azonosítani (Függelék 3).

5.3. Többszörös illesztés és faszerkesztés

A kiválasztott szekvenciákra többszörös illesztést végeztem a ClustalX nevű programmal, amelynek eredményét a törzsfa szerkesztésre alkalmas PHYLIP

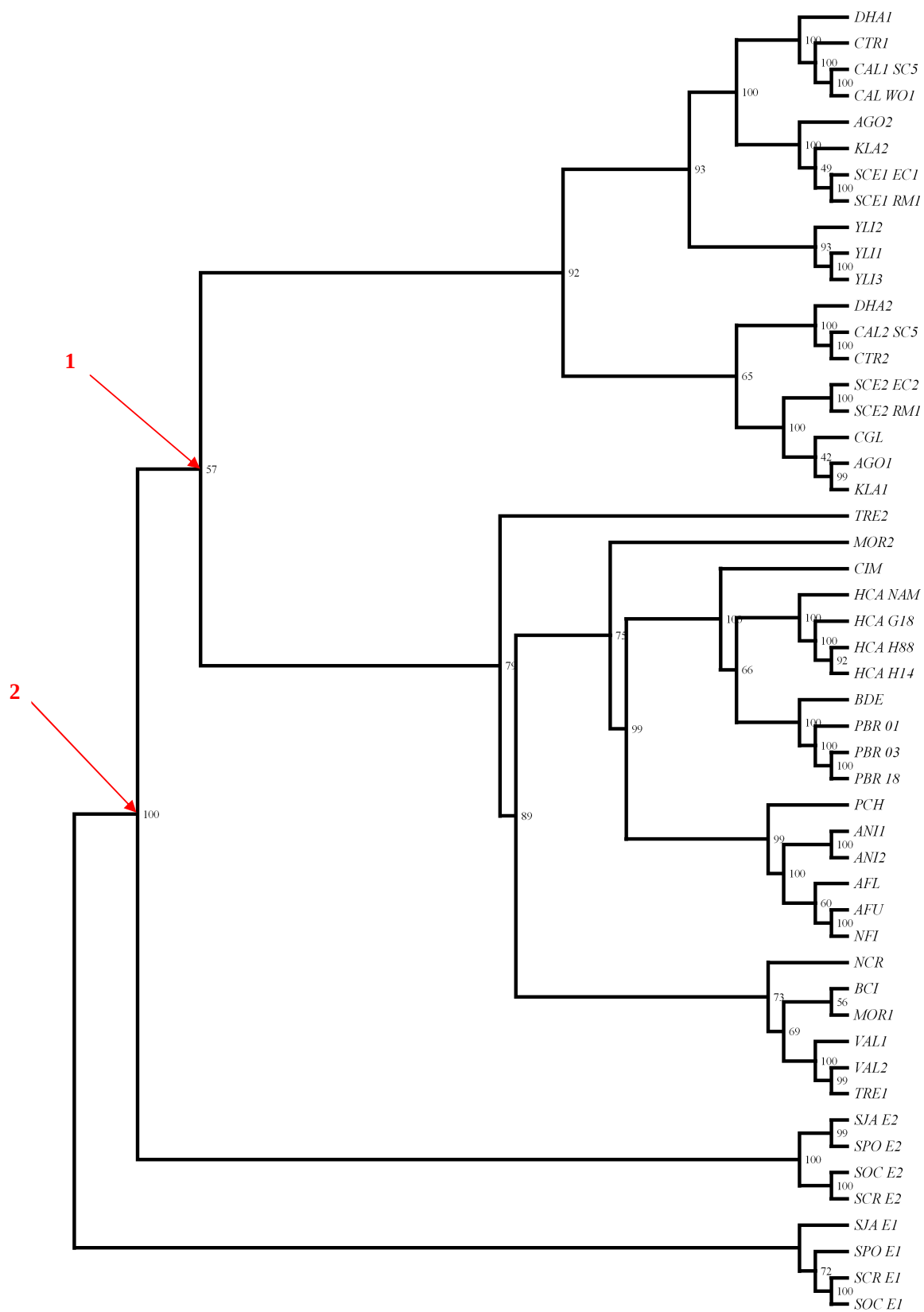
programba vittem át. Törzsfa elkészítéséhez a távolságalapú stratégia (Neighbor joining) módszert választottam, ez a módszer a belső ághosszúságok minimalizálására törekszik, gyakorlatilag a maximális parszimónia módszert alkalmazza távolság adatokra. A Neighbor joining módszer a fában egymás mellé kerülő levelek kiválasztásánál nemcsak a hozzájuk tartozó távolságmátrix értéket, hanem a leveleknek az összes többivel alkotott távolságát is figyelembe veszi. A módosított távolság annál kisebb lesz az eredetinél, minél nagyobb a két levél átlagos távolsága a többitől, mert a nagy átlag (gyors evolúció) megnöveli a két levél relatív közelségét. Eredményül az *eng1* és *eng2* gének lehetséges ortológjait tartalmazó gyökér nélküli fát kaptam. A törzsfa gyökereztetése többféleképpen megvalósítható, tehetjük ezt egy külső csoporttal. Ilyenkor a legjobb a belcsoporthoz legközelebb eső, de abba nem beletartozó fajt alkalmazni. A kapott fán azt az élt, ahol a külső csoport a vizsgált monofiletikus fajokhoz kapcsolódik, a közös őst, illetve a gyökér kijelölésére használjuk. A törzsfa gyökereztetésének egy másik módja az irányított tulajdonságokkal való gyökereztetés, ez akkor lehetséges, ha olyan tulajdonságok alapján végezzük az elemzést, ahol megállapítható, hogy melyik az ősi és melyik a leszármaztatott bélyeg. A filogenetikai (Felsenstein, 1993)ai fát az első módszert szerint szerettem volna gyökereztetni, ezért is tettem kitekintést vizsgálataim során az aszkuszos gombák tagozatából a járomspórás és a bazídiumos gombák irányába, azonban ezen gombafajokban nem találtam megfelelő ortológokat a további analízisekhez. Ezért az elkészült filogenetikai fát a második módszer szerint gyökereztettem. Az *eng1* gén ortológokat tartalmazó *Schizosaccharomyces* csoport jól elkülönült az elsődleges törzsfán, így ezt használtam külső csoportnak a végső törzsfánál (13. ábra).

A 13. ábrán a csomópontokon 100 ismétlésre vonatkoztatott bootstrapértékek, míg az ágak végeinél a vizsgált fajokat szimbolizáló betűkódok láthatóak. A törzsfa ágainak elágazásaiban látható számok az adott elágazás ilyen formában való megtörténének a valószínűségét adják meg, a számok csökkenésével egyre kevésbé lehetünk biztosak az ábrázolás korrektségében. Általánosan elfogadott, hogy 50 alatti bootstrap értéknél az adott elágazás nem fogadható el. Ahogyan az a 13. ábrán is látható, az általam létrehozott törzsfán a bootstrap értékek a legtöbb esetben elegendően nagy valószínűséggel támasztják alá, hogy az evolúció során valóban így történhetett meg a gének elválása az egyes fajokban. A 3. táblázat a 13. ábrához tartozik.

A 13. ábrán az 1-sel jelölt csomópontban feltételezhetően a valódi és a fonalas gombák egy közös őse lehetett, mert a csomópontot követően a két törzs általam vizsgált tagjainak fejlődési vonala teljesen elvált. A 2-sel jelölt csomópontban pedig a *Schizosaccharomyces* nemzetség *eng2* génjének és a többi vizsgált gombafajnak a közös őse található. A filogenetikai fán a *Schizosaccharomyces* nemzetség *eng1* génjének fejlődési vonala egyértelműen elkülönül, mind a nemzetségen belüli *eng2* gén fejlődési ágától, mind pedig az összes többi gombfajban megtalált lehetséges ortológától.

Tehát a *Schizosaccharomyces pombe* citokinézisében szerepet játszó *eng1* és *eng2* gének filogenetikai analizéséből szerkesztett törzsfa alapján valószínűsíthető, hogy először a két paralóg gén evolúciós fejlődése vált el, majd ezt követően az *eng2* génből alakulhatott tovább a fejlődési vonal. Később pedig az *eng2*-ből tovább fejlődött vonalon válhatott el egymástól a fonalas és a valódi gombák világa.

Az *eng1* gént csak a hasadó élesztőkön belül sikerült azonosítanom, aminek lehetséges magyarázata a *Schizosaccharomyces* nemzetségben előforduló egyedi sejtfal szerkezet.



13. ábra: Az *eng1* és *eng2* gének filogenetikai törzsfája (Neighbor-joining fa szerkesztő módszerrel)

3. táblázat: A filogenetika törzsfán használt rövidítések feloldása

Faj	Törzs	Fehérje leírása (NCBI)	Kód
<i>Ashbya gossypii</i>	ATCC 10895	AGL161Cp	AGO1
		AGL208Cp	AGO2
<i>Aspergillus flavus</i>	NRRL3357	Engl1	AFL
<i>Aspergillus fumigatus</i>	Af293	Engl1	AFU
<i>Aspergillus nidulans</i>	FGSC A4	hypothetical protein AN0472.2	ANI1
		putative endo beta 1,3 glucanase	ANI2
<i>Blastomyces dermatitidis</i>	ATCC 18188	glycosyl hydrolase	BDE
<i>Botrytis cinerea</i>	B05.10	hypothetical protein BC1G_00820	BCI
<i>Candida albicans</i>	WO-1	hypothetical protein CAWG_01022	CAL_WO1
	SC5314	hypothetical protein CaO19.10584	CAL1_SC5
		hypothetical protein CaO19.3417	CAL2_SC5
<i>Candida glabrata</i>	CBS 138	hypothetical protein	CGL
<i>Candida tropicalis</i>	MYA-3404	hypothetical protein CTRG_04579	CTR1
		hypothetical protein CTRG_02871	CTR2
<i>Coccidioides immitis</i>	RS	hypothetical protein CIMG_04988	CIM
<i>Debaryomyces hansenii</i>	CBS767	DEHA2B06732p	DHA1
		DEHA2G00726p	DHA2
<i>Histoplasma capsulatum</i>	G186AR	Engl1	HCA_G18
	H88	endo-1,3-beta-glucanase	HCA_H88
	NAm1	hypothetical protein HCAG_06030	HCA_NAM
	H143	Engl1	HCA_H14
<i>Kluyveromyces lactis</i>	NRRL Y-1140	hypothetical protein	KLA1
		hypothetical protein	KLA2
<i>Magnaporthe oryzae</i>	70-15	hypothetical protein MGG_01001	MOR1
		hypothetical protein MGG_05489	MOR2
<i>Neosartorya fischeri</i>	NRRL 181	Engl1	NFI
<i>Neurospora crassa</i>	OR74A	hypothetical protein NCU07076	NCR
<i>Paracoccidioides brasiliensis</i>	Pb01	endo-1,3(4)-beta-glucanase	PBR_01
	Pb03	endo-1,3(4)-beta-glucanase	PBR_03
	Pb18	endo-1,3(4)-beta-glucanase	PBR_18
<i>Penicillium chrysogenum</i>	Wisconsin 54-1255	Pc16g08980	PCH
<i>Saccharomyces cerevisiae</i>	EC1118	Dse4p	SCE1_EC1

A táblázat folytatása a 44. oldalon található

3. táblázat: A filogenetika törzsfa értelmezése

Faj	Törzs	Fehérje leírása (NCBI)	Kód
<i>Saccharomyces cerevisiae</i>	EC1118	Dse4p	SCE1_EC1
		Acf2p	SCE1_EC2
	RM11-1a	conserved hypothetical protein	SCE1_RM1
		conserved hypothetical protein	SCE2_RM1
<i>Schizosaccharomyces cryophilus</i>	NRRL Y-48691T	Eng1	SCR_E1
		Eng2	SCR_E2
<i>Schizosaccharomyces japonicus</i>	yFS275	Eng1	SJA_E1
		Eng2	SJA_E2
<i>Schizosaccharomyces octosporus</i>	yFS28	Eng1	SOC_E1
		Eng2	SOC_E2
<i>Schizosaccharomyces pombe</i>	ATCC 38366/972h ⁻	Eng1	SPO_E1
		Eng2	SPO_E2
<i>Trichoderma reesei</i>	QM6a	glycoside hydrolase family 81	TRE1
		glycoside hydrolase family 81	TRE2
<i>Verticillium albo-atrum</i>	VaMs.102	endo-1,3(4)-beta-glucanase	VAL1
		endo-1,3(4)-beta-glucanase	VAL2
<i>Yarrowia lipolytica</i>	CLIB122	YALI0C12056p	YLI1
		YALI0B17996p	YLI2
		YALI0A07142p	YLI3

6. Összefoglalás

A citokinézis a sejtciklus utolsó lépése, amely során a leánysejtek fizikailag is kettéválnak és két egyedülálló szervezetet alkotnak.

A *S. pombe* két fehérjét tartalmaz a glikozil-hidrozil 81 fehérje családból, amelyek az *eng1* és az *eng2* gének által kódoltak. Mindkét fehérje endo-glükánáz aktivitással rendelkezik, hidrolizálni képesek a β -1,3-glükán láncokat, csak úgy mint ezen fehérje család többi tagja. A két fehérje azonos enzimes aktivitással rendelkezik, azonban a hasadó élesztő életciklusának eltérő részein hatnak. Az Eng1 a lineáris β -1,3-glükant tartalmazó elsődleges szeptumot bontja a sejtosztódás utolsó lépéseként, az *eng2* által kódolt fehérje, pedig más hidrolizáló enzimekkel együtt szükséges az aszkuszfal endolíziséhez a sporuláció utolsó lépésében.

Munkám során a vizsgált 30 gombafaj közül BLASTp analízisek segítségével 26 esetben sikerült *eng1* és *eng2* gén ortológokat azonosítani. A fennmaradó négy esetben az általam külső csoportként használt járomspórás *Phycomyces blakesleeanus*, *Rhizopus oryzae*, valamint a bazídiumos gombák tagozatába tartozó *Cryptococcus neoformans*, *Ustilago maydis* fajok voltak, amelyekben nem találtam meg a hasadó élesztő sejtosztódásában szerepet játszó *eng1* és *eng2* gének ortológjait. Ez pedig arra ad következtetést, hogy ezen gének valószínűleg csak az aszkuszos gombák tagozatába tartozó fajoknál fordulnak elő.

Az azonosított ortológ szekvenciákat annak érdekében, hogy az *eng1* és *eng2* géneket el tudjam különíteni az egyes fajokban, további vizsgálatoknak vetettem alá. Pfam adatbázis segítségével doméneket próbáltam meghatározni az egyes fehérje szekvenciákban, illetve SignalP 3.0 használatával a hasadó élesztőben lévő Eng1-re jellemző szekrécióhoz szükséges szignálpeptid azonosítására tettem kísérletet.

A két gén által kódolt fehérjék szekvenciája közti nagyfokú hasonlóság, és a BLASTp, valamint a fehérje szerkezetet vizsgálatából származó eredmények ellentmondásai miatt, jelen dolgozatomban nem teszek kísérletet a gén ortológok pontos azonosítására. Elképzelhető, hogy a szignálpeptid leesik vagy keletkezik az evolúció során.

Az eredményül kapott filogenetikai törzsfán szerepelnek mind az *eng1*, mind pedig az *eng2* gének lehetséges ortológjai (13. ábra). Valószínűsíthető, hogy a hasadó élesztőben található *eng1* és *eng2* paralóg gének evolúciós fejlődése vált el legkorábban,

majd ezt követően az *eng2* génből alakulhatott tovább a fejlődési vonal. Ennek magyarázata lehet, hogy a hasadó élesztők csoportján belül találkozhatunk csak ehhez hasonló sejtfallal. Az *eng2*-ből tovább fejlődött evolúciós irányon válhatott el egymástól később egy közös őst követően a fonalas és a valódi gombák világa.

További kutatásokat szeretnék végezni a fehérjék szekvenciájában esetlegesen fellelhető domének meghatározásában, illetve a két gén által kódolt fehérje pontos azonosítására, és így a más fajokban megtalálható feltételezett ortológok elkülönítésére. A filogenetikai fa szerkesztésében a gyökereztetést szeretném külső csoporttal megvalósítani, ezért tervezem, hogy egyéb, távolabbi aszakuszos gombákra is kiterjesztem vizsgálataimat. Emellett további gének filogenetikai vizsgálatát is tervezem.

7. Irodalomjegyzék

- Alberts, B., D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. 2009. *Essential Cell Biology*. Garland Publishing, Inc., New York.
- Bathe, M., and F. Chang. 2009. Cytokinesis and the contractile ring in fission yeast: towards a systems-level understanding. *Trends in Microbiology* 18:38-45.
- Bendtsen, J.D., H. Nielsen, G. von Heijne, and S. Brunak. 2004. Improved Prediction of Signal Peptides: SignalP 3.0. *Journal of Molecular Biology* 340:783-795.
- Burgess, D.R., and C. F. 2005. Site selection for the cleavage furrow at cytokinesis. *Trends in Cell Biology* 15:156-162.
- Campbell, A.M., Heyer, Laurie J. 2004. Genomika, proteomika, bioinformatika. Medicina Könyvkiadó Rt., Budapest.
- Dekker, N., D. Speijer, C.H. Grün, M. van den Berg, A. de Haan, and F. Hochstenbach. 2004. Role of the α -Glucanase Agn1p in Fission-Yeast Cell Separation. *Molecular Biology of the Cell* 15:3903-3914.
- del Dedo, J.E., E. Dueñas, Y. Arnáiz, F. del Rey, and C.R. Vázquez de Aldana. 2009. β -Glucanase Eng2 Is Required for Ascus Wall Endolysis after Sporulation in the Fission Yeast *Schizosaccharomyces pombe*. *Eukaryotic Cell* 8:1278-1286.
- Enczi, K., M. Yamaguchi, and M. Sipiczki. 2007. Morphology transition genes in the dimorphic fission yeast *Schizosaccharomyces japonicus*. *Antonie van Leeuwenhoek* 92:143-154.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17:368-376.
- Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. In Distributed by the author. Department of Genetics, University of Washington, Seattle.
- Finn, R.D., J. Mistry, J. Tate, P. Coggill, A. Heger, J.E. Pollington, O.L. Gavin, P. Gunasekaran, G. Ceric, K. Forslund, L. Holm, E.L.L. Sonnhammer, S.R. Eddy, and A. Bateman. 2010. The Pfam protein families database. *Nucleic Acids Research* 38:211-222.
- Forsburg, S.L., and P. Nurse. 1991. Cell-cycle regulation in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Annual Review of Cell and Developmental Biology* 7:227-256.
- Gould, K.L., and V. Simanis. 1997. The control of septum formation in fission yeast. *Genes & Development* 11:2939-2951.

- Helston, R.M., J.A. Box, W. Tang, and P. Baumann. 2010. *Schizosaccharomyces cryophilus* sp. nov., a new species of fission yeast. *FEMS Yeast Research* 10:779–786.
- Jin, G., L. Nakhleh, S. Snir, and T. Tuller. 2007. Inferring Phylogenetic Networks by the Maximum Parsimony Criterion: A Case Study. *Molecular Biology and Evolution* 24:324-337.
- Johnson, B.F., B.Y. Yoo, and G.B. Calleja. 1973. Cell division in yeasts: movement of organelles associated with cell plate growth of *Schizosaccharomyces pombe*. *The Journal of Bacteriology* 115:358-366.
- Kim, D.-U., Jacqueline Hayles, Dongsup Kim, Valerie Wood, Han-Oh Park, Misun Won, Hyang-Sook Yoo, Trevor Duhig, Miyoung Nam, Georgia Palmer, Sangjo Han, Linda Jeffery, Seung-Tae Baek, Hyemi Lee, Young Sam Shim, Minh Lee, Lila Kim, Kyung-Sum Heo, Eun Joo Noh, Ah-Reum Lee, Young-Joo Jang, Kyung-Sook Chung, Shin-Jung Choi, Jo-Young Park, Youngwoo Park, Hwan Mook Kim, Song-Kyu Park, Hae-Joon Park, Eun-Jung Kang, Hyong Bai Kim, Hyun-Sam Kang, Hee-Moon Park, Kyunghoon Kim, Kiwon Song, Kyung Bin Song, Paul Nurse & Kwan-Lae Hoe. 2010. Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. *Nature Biotechnology* 28:617-623.
- MacNeill, S.A., and P.A. Fantes. 1995. Controlling entry into mitosis in fission yeast. In *Cell Cycle Control*. C. Hutchison and D.M. Glover, editors. Oxford University Press, Oxford. 63-105.
- Maller, J.L. 1994. Biochemistry of cell cycle checkpoints at the G2/M and metaphase/anaphase transitions. *Seminars in Cell and Developmental Biology* 5:183-190.
- Martín-Cuadrado, A.B., E. Dueñas, M. Sipiczki, C.R. Vázquez de Aldana, and F. del Rey. 2003. The endo- β -1,3-glucanase eng1p is required for dissolution of the primary septum during cell separation in *Schizosaccharomyces pombe*. *Journal of Cell Science* 116:1689-1698.
- Morgan, D.O. 2007. *The Cell Cycle*. New Science Press Ltd, London, UK.
- Mount, D.W. 2004. *Bioinformatics: Sequence and Genome Analysis*, University of Arizona, Tucson.
- Murray, A. 1994. Cell cycle checkpoints. *Current Opinion in Cell Biology* 6:872-876.
- Nasmyth, K. 2001. A prize for proliferation. *Cell* 107:689-701.
- Nielsen, H., S. Brunak, and G. von Heijne. 1999. Machine learning approaches for the prediction of signal peptides and other protein sorting signals. *Protein Engineering* 12:3-9.

- Rhind, N., Z. Chen, M. Yassour, D.A. Thompson, and *et al.* 2011. Comparative Functional Genomics of the Fission Yeasts. *Science* 332:930-936.
- Ribár, B., A. Bánrévi, and M. Sipiczki. 1997. *sep1⁺* Encodes a transcription-factor homologue of the HNF-3/forkhead DNA-binding-domain family in *Schizosaccharomyces pombe*. *Gene* 202:1-5.
- Saitou, N., and N. Masatoshi. 1987. The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular Biology and Evolution* 4:406-425.
- Sipiczki, M. 1995. Phylogenesis of fission yeast. Contradictions surrounding the origin of a century old genus. *Antonie van Leeuwenhoek* 68:119-149.
- Sipiczki, M. 2000. Where does fission yeast sit on the tree of life? *Genome Biology* 1:1011.1-1011.4.
- Sipiczki, M. 2007. Splitting of the fission yeast septum. *FEMS Yeast Research* 7:761-770.
- Sipiczki, M., K. Takeo, M. Yarnaguchi, S. Yoshida, and I. Miklos. 1998. Environmentally controlled dimorphic cycle in a fission yeast. *Microbiology* 144:1319-1330.
- Sveiczzer, A., J. Tyson, and B. Novak. 2004. Modelling the fission yeast cell cycle. *Briefings in Functional Genomics and Proteomics* 2:298-307.
- Wang, H., Z. Xu, L. Gao, and B. Hao. 2009. A fungal phylogeny based on 82 complete genomes using the composition vector method. *BMC Evolutionary Biology* 9.
- Wood, V., R. Gwilliam, M.-A. Rajandream, M. Lyne, R. Lyne, A. Stewart, J. Sgouros, N. Peat, J. Hayles, S. Baker, D. Basham, and S. Bowman. 2002. The genome sequence of *Schizosaccharomyces pombe*. *Nature* 415:871-880.
- Zetterberg, A., O. Larsson, and K.G. Wiman. 1995. What is the restriction point? *Current Opinion in Cell Biology* 7:835-842.

Faj	<i>S. pombe</i>		<i>S. japonicus</i>		<i>S. octosporus</i>		<i>N. crassa</i>
Fehérje leírása (NCBI)	Eng1	Eng2	Eng1	endo-1,3(4)-beta-glucanase	eng1	eng2	hypothetical protein NCU07076
kód	NP_593162.1	NP_594547.1	SJAG_00667.4	XP_002172286.1	SOCG_02138	SOCG_05926.5	XP_960341.1
Fehérje hossza (aa)	1016	706	1008	681	1043	707	912
pos		417/699 (60%)	733/1014 (72%)	393/688 (57%)	746/1037 (72%)	415/699 (59%)	392/729 (54%)
ident		287/699 (41%)	578/1014 (57%)	290/688 (42%)	591/1037 (57%)	286/699 (41%)	268/729 (37%)
gaps		24/699 (3%)	14/1014 (1%)	34/688 (5%)	48/1037 (5%)	22/699 (3%)	60/729 (8%)
score		513 bits (1320)	1098 bits (2840)	516 bits (1329)	1189 bits (3077)	528 bits (1359)	451 bits (1160)
E		2,00E-149	0.0	2,00E-150	0.0	7,00E-154	1,00E-130
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1	NP_594547.1	NP_593162.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng1	Eng2	Eng1	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1	NP_593162.1	NP_594547.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng2	Eng1	Eng2	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>S. cryophilus</i>		<i>Neosartorya fischeri</i>	<i>Ashbya gossypii</i>		<i>Botrytis cinerea</i>
Törzs	972h ⁻	NRRL Y-48691T	NRRL Y-48691T	NRRL 181	ATCC 10895	ATCC 10895	B05.10
Fehérje leírása (NCBI)	Eng1	eng1	eng2	Eng1	AGL208Cp	AGL161Cp	hypothetical protein BC1G_00820
kód	NP_593162.1	SPOG_00688.3	SPOG_02717.3	XP_001265240.1	NP_986459.2		XP_001560792.1
Fehérje hossza (aa)	1016	1055	706	721	861	770	410
pos		750/1057 (71%)	413/710 (58%)	411/699 (59%)	410/714 (57%)	395/724 (55%)	197/393 (50%)
ident		593/1057 (56%)	290/710 (41%)	296/699 (42%)	271/714 (38%)	276/724 (38%)	131/393 (33%)
gaps		51/1057 (5%)	24/710 (3%)	35/699 (5%)	36/714 (5%)	36/724 (5%)	22/393 (6%)
score		1186 bits (3068)	524 bits (1349)	519 bits (1336)	471 bits (1213)	473 bits (1216)	183 bits (464)
E		0.0	1,00E-152	3,00E-151	8,00E-154	3,00E-137	2,00E-50
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_593162.1	NP_594547.1	NP_594547.1	NP_593162.1	NP_594547.1	NP_594547.1
	leírás	Eng1	Eng2	Eng2	Eng1	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_594547.1	NP_593162.1	NP_593162.1	NP_594547.1	NP_593162.1	NP_593162.1
	leírás	Eng2	Eng1	Eng1	Eng2	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Debaryomyces hansenii</i>		<i>Verticillium albo-atrum</i>		<i>Coccidioides immitis</i>
Törzs	972h ⁻	CBS767	CBS767	VaMs.102	VaMs.102	RS
Fehérje leírása (NCBI)	Eng1	DEHA2B06732p	DEHA2G00726p	endo-1,3(4)-beta-glucanase	endo-1,3(4)-beta-glucanase	hypothetical protein CIMG_04988
kód	NP_593162.1	XP_457251.2	XP_461565.2	XP_003005468.1	XP_003001779.1	XP_001245547.1
Fehérje hossza (aa)	1016	1159	773	838	840	720
pos		435/734 (59%)	392/727 (54%)	403/702 (57%)	219/321 (68%)	405/720 (56%)
ident		297/734 (40%)	268/727 (37%)	286/702 (41%)	162/321 (50%)	296/720 (41%)
gaps		37/734 (5%)	40/727 (6%)	29/702 (4%)	5/321 (2%)	38/720 (5%)
score		513 bits (1322)	444 bits (1141)	499 bits (1285)	350 bits (897)	501 bits (1289)
E		3,00E-149	2,00E-128	3,00E-145	3,00E-100	1,00E-145
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_593162.1	NP_594547.1	NP_594547.1	NP_593162.1	NP_594547.1
	leírás	Eng1	Eng2	Eng2	Eng1	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_594547.1	NP_593162.1	NP_593162.1	NP_594547.1	NP_593162.1
	leírás	Eng2	Eng1	Eng1	Eng2	Eng1

Faj	<i>S. pombe</i>	<i>Aspergillus fumigatus</i>	<i>Aspergillus flavus</i>	<i>Aspergillus nidulans</i>		<i>Penicillium chrysogenum</i>
Törzs	972h ⁻	Af293	NRRL3357	FGSC A4	FGSC A4	Wisconsin 54-1255
Fehérje leírása (NCBI)	Eng1	Eng1	Eng1	hypothetical protein AN0472.2	putative endo beta 1,3 glucanase	Pc16g08980
kód	NP_593162.1	XP_750173.1	XP_002374405.1	XP_658076.1	CBF89420.1	XP_002561217.1
Fehérje hossza (aa)	1016	974	910	934	907	894
pos		405/706 (57%)	394/691 (57%)	407/707 (58%)	401/701 (57%)	391/719 (54%)
ident		294/706 (42%)	280/691 (41%)	297/707 (42%)	292/701 (42%)	288/719 (40%)
gaps		41/706 (6%)	31/691 (4%)	35/707 (5%)	35/701 (5%)	31/719 (4%)
score		514 bits (1325)	504 bits (1299)	528 bits (1359)	518 bits (1333)	491 bits (1264)
E		8,00E-150	8,00E-147	1,00E-153	1,00E-150	8,00E-143
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng2	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng1	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Candida tropicalis</i>		<i>Candida glabrata</i>	<i>Yarrowia lipolytica</i>		
Törzs	972h ⁻	MYA-3404	MYA-3404	CBS 138	CLIB122	CLIB122	CLIB122
Fehérje leírása (NCBI)	Eng1	hypothetical protein CTRG_04579	hypothetical protein CTRG_02871	hypothetical protein	YALIO0C12056p	YALIO0B17996p	YALIO0A07142p
kód	NP_593162.1	XP_002550281.1	XP_002548574.1	XP_449546.1	XP_501746.2	XP_501038.1	XP_499836.2
Fehérje hossza (aa)	1016	1156	766	770	763	935	751
pos		416/721 (58%)	373/693 (54%)	392/715 (55%)	415/742 (56%)	407/723 (56%)	404/711 (57%)
ident		283/721 (39%)	246/693 (35%)	279/715 (39%)	289/742 (39%)	290/723 (40%)	285/711 (40%)
gaps		48/721 (7%)	34/693 (5%)	44/715 (6%)	37/742 (5%)	48/723 (7%)	31/711 (4%)
score		468 bits (1203)	407 bits (1045)	459 bits (1182)	504 bits (1297)	507 bits (1306)	494 bits (1273)
E		1,00E-135	2,00E-117	2,00E-133	1,00E-146	1,00E-147	6,00E-144
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_593162.1	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng1	Eng2	Eng2	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_594547.1	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng2	Eng1	Eng1	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Candida albicans</i>			<i>Magnaporthe oryzae</i>		<i>Blastomyces dermatitidis</i>
Törzs	972h ⁻	WO-1	SC5314	SC5314	70-15	70-15	ATCC 18188
Fehérje leírása (NCBI)	Eng1	hypothetical protein CAWG_01022	hypothetical protein CaO19.10584	hypothetical protein CaO19.3417	hypothetical protein MGG_01001	hypothetical protein MGG_05489	glycosyl hydrolase
kód	NP_593162.1	EEQ42800.1	XP_721293.1	XP_716726.1	XP_368243.2	XP_360114.2	EGE85606.1
Fehérje hossza (aa)	1016	1143	1146	734	820	811	932
pos		410/714 (57%)	410/714 (57%)	386/703 (55%)	339/638 (53%)	365/707 (52%)	400/721 (55%)
ident		275/714 (39%)	275/714 (39%)	256/703 (36%)	247/638 (39%)	258/707 (36%)	293/721 (41%)
gaps		34/714 (5%)	34/714 (5%)	29/703 (4%)	34/638 (5%)	43/707 (6%)	36/721 (5%)
score		483 bits (1244)	482 bits (1241)	447 bits (1149)	411 bits (1057)	409 bits (1052)	504 bits (1299)
E		3,00E-140	6,00E-140	1,00E-129	7,00E-119	3,00E-118	8,00E-147
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_593162.1	NP_593162.1	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng1	Eng1	Eng2	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_594547.1	NP_594547.1	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng2	Eng2	Eng1	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Paracoccidioides brasiliensis</i>			<i>Trichoderma reesei</i>	
Törzs	972h ⁻	Pb01	Pb03	Pb18	QM6a	QM6a
Fehérje leírása (NCBI)	Eng1	endo-1,3(4)-beta-glucanase	endo-1,3(4)-beta-glucanase	endo-1,3(4)-beta-glucanase	glycoside hydrolase family 81	glycoside hydrolase family 81
kód	NP_593162.1	XP_002797710.1	EEH16341.1	EEH43059.1	EGR47329.1	EGR49603.1
Fehérje hossza (aa)	1016	782	845	592	712	863
pos		394/687 (57%)	298/480 (62%)	201/383 (52%)	392/714 (55%)	212/479 (44%)
ident		299/687 (44%)	232/480 (48%)	136/383 (36%)	268/714 (38%)	138/479 (29%)
gaps		34/687 (5%)	18/480 (4%)	18/383 (5%)	29/714 (4%)	43/479 (9%)
score		519 bits (1336)	437 bits (1123)	224 bits (571)	476 bits (1226)	179 bits (453)
E		4,00E-151	2,00E-126	1,00E-62	2,00E-138	1,00E-48
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1	NP_594547.1	NP_594547.1	NP_593162.1
	leírás	Eng2	Eng1	Eng2	Eng2	Eng1
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1	NP_593162.1	NP_593162.1	NP_594547.1
	leírás	Eng1	Eng2	Eng1	Eng1	Eng2

Faj	<i>S. pombe</i>	<i>Saccharomyces cerevisiae</i>			
Törzs	972h ⁻	EC1118	RM11-1a	RM11-1a	EC1118
Fehérje leírása (NCBI)	Eng1	Dse4p	conserved hypothetical protein	conserved hypothetical protein	Acf2p
kód	NP_593162.1	CAY82261.1	EDV12201.1	EDV09439.1	CAY81379.1
Fehérje hossza (aa)	1016	1117	1117	779	779
pos		425/723 (59%)	424/723 (59%)	382/714 (54%)	381/714 (53%)
ident		286/723 (40%)	286/723 (40%)	266/714 (37%)	265/714 (37%)
gaps		42/723 (6%)	42/723 (6%)	31/714 (4%)	31/714 (4%)
score		478 bits (1229)	477 bits (1228)	442 bits (1138)	441 bits (1134)
E		1,00E-138	2,00E-138	3,00E-128	9,00E-128
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Histoplasma capsulatum</i>			
Törzs	972h ⁻	G186AR	H88	NAm1	H143
Fehérje leírása (NCBI)	Eng1	Eng1	endo-1,3-beta-glucanase	hypothetical protein HCAG_06030	Eng1
kód	NP_593162.1	EEH05945.1	EGC49676.1	XP_001538425.1	EER41848.1
Fehérje hossza (aa)	1016	848	848	839	797
pos		404/724 (56%)	401/721 (56%)	393/717 (55%)	376/720 (52%)
ident		293/724 (40%)	287/721 (40%)	287/717 (40%)	270/720 (38%)
gaps		41/724 (6%)	35/721 (5%)	41/717 (6%)	84/720 (12%)
score		500 bits (1287)	489 bits (1260)	479 bits (1234)	439 bits (1129)
E		2,00E-145	3,00E-142	3,00E-139	3,00E-127
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Kluyveromyces lactis</i>	
Törzs	972h ⁻	NRRL Y-1140	NRRL Y-1140
Fehérje leírása (NCBI)	Eng1	hypothetical protein	hypothetical protein
kód	NP_593162.1	XP_453814.1	XP_455611.1
Fehérje hossza (aa)	1016	806	1013
pos		389/711 (55%)	398/721 (55%)
ident		278/711 (39%)	269/721 (37%)
gaps		32/711 (5%)	46/721 (6%)
score		482 bits (1241)	451 bits (1159)
E		4,00E-140	2,00E-130
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1
	leírás	Eng2	Eng1
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1
	leírás	Eng1	Eng2

Faj	<i>S. pombe</i>		<i>S. japonicus</i>		<i>S. octosporus</i>		<i>N. crassa</i>
Fehérje leírása (NCBI)	Eng1	Eng2	Eng1	endo-1,3(4)-beta-glucanase	eng1	eng2	hypothetical protein NCU07076
kód	NP_593162.1	NP_594547.1	SJAG_00667.4	XP_002172286.1	SOCG_02138	SOCG_05926.5	XP_960341.1
Fehérje hossza (aa)	1016	706	1008	681	1043	707	912
pos		417/699 (60%)	733/1014 (72%)	393/688 (57%)	746/1037 (72%)	415/699 (59%)	392/729 (54%)
ident		287/699 (41%)	578/1014 (57%)	290/688 (42%)	591/1037 (57%)	286/699 (41%)	268/729 (37%)
gaps		24/699 (3%)	14/1014 (1%)	34/688 (5%)	48/1037 (5%)	22/699 (3%)	60/729 (8%)
score		513 bits (1320)	1098 bits (2840)	516 bits (1329)	1189 bits (3077)	528 bits (1359)	451 bits (1160)
E		2,00E-149	0.0	2,00E-150	0.0	7,00E-154	1,00E-130
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1	NP_594547.1	NP_593162.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng1	Eng2	Eng1	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1	NP_593162.1	NP_594547.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng2	Eng1	Eng2	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>S. cryophilus</i>		<i>Neosartorya fischeri</i>	<i>Ashbya gossypii</i>		<i>Botrytis cinerea</i>
Törzs	972h ⁻	NRRL Y-48691T	NRRL Y-48691T	NRRL 181	ATCC 10895	ATCC 10895	B05.10
Fehérje leírása (NCBI)	Eng1	eng1	eng2	Eng1	AGL208Cp	AGL161Cp	hypothetical protein BC1G_00820
kód	NP_593162.1	SPOG_00688.3	SPOG_02717.3	XP_001265240.1	NP_986459.2		XP_001560792.1
Fehérje hossza (aa)	1016	1055	706	721	861	770	410
pos		750/1057 (71%)	413/710 (58%)	411/699 (59%)	410/714 (57%)	395/724 (55%)	197/393 (50%)
ident		593/1057 (56%)	290/710 (41%)	296/699 (42%)	271/714 (38%)	276/724 (38%)	131/393 (33%)
gaps		51/1057 (5%)	24/710 (3%)	35/699 (5%)	36/714 (5%)	36/724 (5%)	22/393 (6%)
score		1186 bits (3068)	524 bits (1349)	519 bits (1336)	471 bits (1213)	473 bits (1216)	183 bits (464)
E		0.0	1,00E-152	3,00E-151	8,00E-154	3,00E-137	2,00E-50
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_593162.1	NP_594547.1	NP_594547.1	NP_593162.1	NP_594547.1	NP_594547.1
	leírás	Eng1	Eng2	Eng2	Eng1	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_594547.1	NP_593162.1	NP_593162.1	NP_594547.1	NP_593162.1	NP_593162.1
	leírás	Eng2	Eng1	Eng1	Eng2	Eng1	Eng1

Faj	<i>S. pombe</i>		<i>S. japonicus</i>		<i>S. octosporus</i>		<i>Neurospora crassa</i>
Törzs	972h ⁻	972h ⁻	yFS275	yFS275	yFS28	yFS28	OR74A
Fehérje leírása (NCBI)	Eng2	Eng1	Eng2	endo-1,3(4)-beta-glucanase	Eng2	Eng1	hypothetical protein NCU07076
kód	NP_594547.1	NP_593162.1	SJAG_01017.4	XP_002171938.1	SOCG_05926.5	SOCG_02138.5	XP_960341.1
Fehérje hossza (aa)	706	1016	682	1007	707	1043	912
pos	417/699 (60%)		528/707 (75%)	424/707 (60%)	514/706 (73%)	414/695 (60%)	403/726 (56%)
ident	287/699 (41%)		441/707 (62%)	282/707 (40%)	416/706 (59%)	290/695 (42%)	268/726 (37%)
gaps	24/699 (3%)		30/707 (4%)	23/707 (3%)	2/706 (0%)	21/695 (3%)	55/726 (8%)
score	520 bits (1338)		908 bits (2347)	500 bits (1287)	876 bits (2263)	531 bits (1367)	451 bits (1159)
E	2,00E-151		0.0	1,00E-145	0.0	8,00E-155	1,00E-130
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_593162.1	NP_594547.1	NP_593162.1	NP_594547.1	NP_593162.1	NP_594547.1
	leírás	Eng1	Eng2	Eng1	Eng2	Eng1	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_594547.1	NP_593162.1	NP_594547.1	NP_593162.1	NP_594547.1	NP_593162.1
	leírás	Eng2	Eng1	Eng2	Eng1	Eng2	Eng1

Faj	<i>S. pombe</i>	<i>S. cryophilus</i>		<i>Neosartorya fischeri</i>	<i>Ashbya gossypii</i>		<i>Botrytis cinerea</i>
Törzs	972h ⁻	NRRL Y-48691T	NRRL Y-48691T	NRRL 181	ATCC 10895	ATCC 10895	B05.10
Fehérje leírása (NCBI)	Eng2	Eng2	Eng1	Eng1	AGL161Cp	AGL208Cp	hypothetical protein BC1G_00820
kód	NP_594547.1	SPOG_02717.3	SPOG_00688.3	XP_001265240.1	NP_986506.1	NP_986459.2	XP_001560792.1
Fehérje hossza (aa)	706	706	1055	721	770	861	410
pos	499/707 (71%)		411/694 (59%)	421/719 (59%)	381/689 (55%)	389/722 (54%)	217/392 (55%)
ident	405/707 (57%)		285/694 (41%)	302/719 (42%)	279/689 (40%)	263/722 (36%)	142/392 (36%)
gaps	3/707 (0%)		20/694 (3%)	26/719 (4%)	44/689 (6%)	32/722 (4%)	16/392 (4%)
score	850 bits (2196)		522 bits (1344)	554 bits (1428)	470 bits (1210)	422 bits (1085)	230 bits (586)
E	0.0		4,00E-152	5,00E-162	9,00E-137	3,00E-122	1,00E-64
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1	NP_594547.1	NP_594547.1	NP_593162.1	NP_594547.1
	leírás	Eng2	Eng1	Eng2	Eng2	Eng1	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1	NP_593162.1	NP_593162.1	NP_594547.1	NP_593162.1
	leírás	Eng1	Eng2	Eng1	Eng1	Eng2	Eng1

Faj	<i>S. pombe</i>	<i>Debaryomyces hansenii</i>		<i>Verticillium albo-atrum</i>		<i>Coccidioides immitis</i>
Törzs	972h ⁻	CBS767	CBS767	VaMs.102	VaMs.102	RS
Fehérje leírása (NCBI)	Eng2	DEHA2G00726p	DEHA2B06732p	endo-1,3(4)-beta-glucanase	endo-1,3(4)-beta-glucanase	hypothetical protein CIMG_04988
kód	NP_594547.1	XP_461565.2	XP_457251.2	XP_003005468.1	XP_003001779.1	XP_001245547.1
Fehérje hossza (aa)	706	773	1159	838	840	720
pos		414/724 (57%)	421/729 (58%)	424/713 (59%)	202/325 (62%)	403/724 (56%)
ident		283/724 (39%)	295/729 (40%)	292/713 (41%)	157/325 (48%)	300/724 (41%)
gaps		40/724 (6%)	40/729 (5%)	22/713 (3%)	4/325 (1%)	27/724 (4%)
score		508 bits (1307)	496 bits (1277)	548 bits (1411)	310 bits (795)	536 bits (1381)
E		6,00E-148	3,00E-144	5,00E-160	1,00E-88	1,00E-156
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1	NP_594547.1	NP_593162.1	NP_594547.1
	leírás	Eng2	Eng1	Eng2	Eng1	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1	NP_593162.1	NP_594547.1	NP_593162.1
	leírás	Eng1	Eng2	Eng1	Eng2	Eng1

Faj	<i>S. pombe</i>	<i>Aspergillus fumigatus</i>	<i>Aspergillus flavus</i>	<i>Aspergillus nidulans</i>		<i>Penicillium chrysogenum</i>
Törzs	972h ⁻	Af293	NRRL3357	FGSC A4	FGSC A4	Wisconsin 54-1255
Fehérje leírása (NCBI)	Eng2	Engl1	Engl1	hypothetical protein AN0472.2	putative endo beta 1,3 glucanase	Pc16g08980
kód	NP_594547.1	XP_750173.1	XP_002374405.1	XP_658076.1	CBF89420.1	XP_002561217.1
Fehérje hossza (aa)	706	974	910	934	907	894
pos		419/728 (58%)	402/713 (56%)	414/718 (58%)	409/712 (57%)	388/719 (54%)
ident		306/728 (42%)	294/713 (41%)	299/718 (42%)	294/712 (41%)	278/719 (39%)
gaps		34/728 (5%)	24/713 (3%)	24/718 (3%)	24/712 (3%)	26/719 (4%)
score		552 bits (1423)	535 bits (1377)	544 bits (1401)	536 bits (1380)	494 bits (1271)
E		2,00E-161	4,00E-156	9,00E-159	2,00E-156	9,00E-144
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng2	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng1	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Candida tropicalis</i>		<i>Candida glabrata</i>	<i>Yarrowia lipolytica</i>		
Törzs	972h ⁻	MYA-3404	MYA-3404	CBS 138	CLIB122	CLIB122	CLIB122
Fehérje leírása (NCBI)	Eng2	hypothetical protein CTRG_02871	hypothetical protein CTRG_04579	hypothetical protein	YALI0C12056p	YALI0B17996p	YALI0A07142p
kód	NP_594547.1	XP_002548574.1	XP_002550281.1	XP_449546.1	XP_501746.2	XP_501038.1	XP_499836.2
Fehérje hossza (aa)	706	766	1156	770	763	935	751
pos		401/715 (56%)	393/722 (54%)	396/708 (56%)	404/715 (57%)	399/684 (58%)	404/720 (56%)
ident		282/715 (39%)	274/722 (38%)	277/708 (39%)	280/715 (39%)	273/684 (40%)	285/720 (40%)
gaps		47/715 (7%)	32/722 (4%)	36/708 (5%)	26/715 (4%)	25/684 (4%)	29/720 (4%)
score		481 bits (1238)	433 bits (1113)	491 bits (1263)	525 bits (1352)	500 bits (1288)	492 bits (1266)
E		5,00E-140	2,00E-125	7,00E-143	3,00E-153	1,00E-145	3,00E-143
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng1	Eng2	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng2	Eng1	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Candida albicans</i>			<i>Magnaporthe grisea</i>		<i>Blastomyces dermatitidis</i>
Törzs	972h ⁻	SC5314	WO-1	SC5314	70-15	70-15	ATCC 18188
Fehérje leírása (NCBI)	Eng2	hypothetical protein CaO19.3417	hypothetical protein CAWG_01022	hypothetical protein CaO19.10584	hypothetical protein MGG_01001	hypothetical protein MGG_05489	glycosyl hydrolase
kód	NP_594547.1	XP_716726.1	EEQ42800.1	XP_721293.1	XP_368243.2	XP_360114.2	EGE85606.1
Fehérje hossza (aa)	706	734	1143	1146	820	811	932
pos		404/709 (57%)	399/725 (55%)	392/701 (56%)	361/640 (56%)	379/722 (52%)	399/727 (55%)
ident		288/709 (41%)	287/725 (40%)	280/701 (40%)	259/640 (40%)	260/722 (36%)	294/727 (40%)
gaps		43/709 (6%)	36/725 (5%)	35/701 (5%)	25/640 (4%)	36/722 (5%)	29/727 (4%)
score		498 bits (1282)	454 bits (1169)	452 bits (1163)	438 bits (1126)	429 bits (1104)	525 bits (1353)
E		4,00E-145	7,00E-132	4,00E-131	6,00E-127	2,00E-124	3,00E-153
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1	NP_593162.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng1	Eng1	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1	NP_594547.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng2	Eng2	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Paracoccidioides brasiliensis</i>			<i>Trichoderma reesei</i>	
Törzs	972h ⁻	Pb01	Pb03	Pb18	QM6a	QM6a
Fehérje leírása (NCBI)	Eng2	endo-1,3(4)-beta-glucanase	endo-1,3(4)-beta-glucanase	endo-1,3(4)-beta-glucanase	glycoside hydrolase family 81	glycoside hydrolase family 81
kód	NP_594547.1	XP_002797710.1	EEH16341.1	EEH43059.1	EGR47329.1	EGR49603.1
Fehérje hossza (aa)	706	782	845	592	712	863
pos		404/695 (58%)	276/477 (58%)	201/383 (52%)	405/712 (57%)	251/562 (45%)
ident		296/695 (43%)	211/477 (44%)	136/383 (36%)	281/712 (39%)	154/562 (27%)
gaps		30/695 (4%)	14/477 (3%)	18/383 (5%)	24/712 (3%)	53/562 (9%)
score		553 bits (1424)	394 bits (1013)	224 bits (571)	533 bits (1373)	180 bits (456)
E		2,00E-161	8,00E-114	1,00E-62	1,00E-155	3,00E-49
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1	NP_594547.1	NP_594547.1	NP_593162.1
	leírás	Eng2	Eng1	Eng2	Eng2	Eng1
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1	NP_593162.1	NP_593162.1	NP_594547.1
	leírás	Eng1	Eng2	Eng1	Eng1	Eng2

Faj	<i>S. pombe</i>	<i>Saccharomyces cerevisiae</i>			
Törzs	972h ⁻	RM11-1a	EC1118	RM11-1a	EC1118
Fehérje leírása (NCBI)	Eng2	conserved hypothetical protein	Dse4p	conserved hypothetical protein	Acf2p
kód	NP_594547.1	EDV12201.1	CAY82261.1	EDV09439.1	CAY81379.1
Fehérje hossza (aa)	706	1117	1117	779	779
pos		408/709 (58%)	409/709 (58%)	393/707 (56%)	392/707 (55%)
ident		279/709 (39%)	279/709 (39%)	277/707 (39%)	276/707 (39%)
gaps		25/709 (4%)	25/709 (4%)	39/707 (6%)	39/707 (6%)
score		486 bits (1252)	486 bits (1252)	474 bits (1220)	473 bits (1216)
E		2,00E-141	2,00E-141	6,00E-138	2,00E-137
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Histoplasma capsulatum</i>			
Törzs	972h ⁻	G186AR	H88	NAm1	H143
Fehérje leírása (NCBI)	Eng2	Eng1	endo-1,3-beta-glucanase	hypothetical protein HCAG_06030	Eng1
kód	NP_594547.1	EEH05945.1	EGC49676.1	XP_001538425.1	EER41848.1
Fehérje hossza (aa)	706	848	848	839	797
pos		405/724 (56%)	406/724 (56%)	398/712 (56%)	383/721 (53%)
ident		299/724 (41%)	298/724 (41%)	288/712 (40%)	283/721 (39%)
gaps		32/724 (4%)	32/724 (4%)	30/712 (4%)	77/721 (11%)
score		533 bits (1373)	528 bits (1359)	521 bits (1342)	479 bits (1233)
E		1,00E-155	6,00E-154	5,00E-152	2,00E-139
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_594547.1	NP_594547.1	NP_594547.1
	leírás	Eng2	Eng2	Eng2	Eng2
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_593162.1	NP_593162.1	NP_593162.1
	leírás	Eng1	Eng1	Eng1	Eng1

Faj	<i>S. pombe</i>	<i>Kluyveromyces lactis</i>	
Törzs	972h ⁻	NRRL Y-1140	NRRL Y-1140
Fehérje leírása (NCBI)	Eng2	hypothetical protein	hypothetical protein
kód	NP_594547.1	XP_453814.1	XP_455611.1
Fehérje hossza (aa)	706	806	1013
pos		409/706 (58%)	395/727 (54%)
ident		303/706 (43%)	272/727 (37%)
gaps		36/706 (5%)	40/727 (6%)
score		507 bits (1306)	434 bits (1117)
E		8,00E-148	8,00E-126
reciprok BLASTp (<i>S.pombe</i>) első találat	kód	NP_594547.1	NP_593162.1
	leírás	Eng2	Eng1
reciprok BLASTp (<i>S.pombe</i>) második találat	kód	NP_593162.1	NP_594547.1
	leírás	Eng1	Eng2

Faj		<i>S. pombe</i>		<i>S. japonicus</i>		<i>S. octosporus</i>		<i>S. cryophilus</i>	
Törzs		972h ⁻	972h ⁻	yFS275	yFS275	yFS28	yFS28	NRRL Y-48691T	NRRL Y-48691T
Kód		NP_593162.1	NP_594547.1	SJAG_00667.4	XP_002172286.1	SOCG_02138	SOCG_05926.5	SPOG_00688.3	SPOG_02717.3
Fehérje hossza (aa)		1016	706	1008	681	1043	707	1055	706
reciprok BLASTp első találat		Eng1	Eng2	Eng1	Eng2	Eng1	Eng2	Eng1	Eng2
NN	Szignálpeptid?	IGEN	NEM	IGEN	heterogén	IGEN	NEM	IGEN	heterogén
HMM	Szignálpeptid valószínűsége	0.990	0.005	0.999	0.004	0.986	0.012	0.993	0.001
	Szignál horgony valószínűsége	0.006	0.000	0.000	0.000	0.009	0.000	0.006	0.000
	Max. hasítási hely valószínűség	0.815	0.003	0.918	0.004	0.770	0.007	0.949	0.001
	Hasítási hely	20 and 21	20 and 21	21 and 22	17 and 18	22 and 23	15 and 16	21 and 22	16 and 17
	Előrejelzés	Szignálpeptid	Nincs szignálpeptid	Szignálpeptid	Nincs szignálpeptid	Szignálpeptid	Nincs szignálpeptid	Szignálpeptid	Nincs szignálpeptid
Heterogán NN	max. C				Nem				Igen
	max. Y				Igen				Nem
	max. S				Igen				Nem
	mean S				Igen				Nem
	D				Igen				Nem
Pfam		Glikozil hidroláz 81, 3 db szénhidrát kötő domén	Glikozil hidroláz 81	Glikozil hidroláz 81, 3 db szénhidrát kötő domén	Glikozil hidroláz 81	Glikozil hidroláz 81, 3 db szénhidrát kötő domén	Glikozil hidroláz 81	Glikozil hidroláz 81, 3 db szénhidrát kötő domén	Glikozil hidroláz 81
Faj		<i>Ashbya gossypii</i>		<i>Kluyveromyces lactis</i>		<i>Debaryomyces hansenii</i>		<i>Verticillium albo-atrum</i>	
Törzs		ATCC 10895	ATCC 10895	NRRL Y-1140	NRRL Y-1140	CBS767	CBS767	VaMs.102	VaMs.102
Kód		NP_986506.1	NP_986459.2	XP_453814.1	XP_455611.1	XP_457251.2	XP_461565.2	XP_003005468.1	XP_003001779.1
Fehérje hossza (aa)		773	861	806	1013	1159	773	838	840
reciprok BLASTp első találat			Eng1	Eng2	Eng1	Eng1	Eng2	Eng2	Eng1
NN	Szignálpeptid?		heterogén	NEM	IGEN	IGEN	NEM	IGEN	heterogén
HMM	Szignálpeptid valószínűsége		0.499	0.000	1.000	0.993	0.000	0.998	0.760
	Szignál horgony valószínűsége		0.000	0.000	0.000	0.007	0.000	0.000	0.000
	Max. hasítási hely valószínűség		0.000	0.000	0.952	0.886	0.000	0.689	0.238
	Hasítási hely		15 and 16	22 and 23	-1 and 0	20 and 21	25 and 26	-1 and 0	16 and 17
	Előrejelzés		Nincs szignálpeptid	Nincs szignálpeptid	Nincs szignálpeptid	Szignálpeptid	Szignálpeptid	Nincs szignálpeptid	Szignálpeptid
Heterogán NN	max. C		Nem						Nem
	max. Y		Nem						Igen
	max. S		Nem						Igen
	mean S		Igen						Igen
	D		Nem						Igen
Pfam		Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81

		<i>Candida glabrata</i>	<i>Candida tropicalis</i>		<i>Coccidioides immitis</i>	<i>Botrytis cinerea</i>
Törzs		CBS 138	MYA-3404	MYA-3404	RS	B05.10
Kód		XP_449546.1	XP_002550281.1	XP_002548574.1	XP_001245547.1	XP_001560792.1
Fehérje hossza (aa)		770	1156	766	720	410
reciprok BLASTp első találat		Eng2	Eng1	Eng2	Eng2	Eng2
NN	Szignálpeptid?	NEM	IGEN	NEM	NEM	NEM
HMM	Szignálpeptid valószínűsége	0.000	0.993	0.000	0.000	0.000
	Szignál horgony valószínűsége	0.000	0.000	0.000	0.000	0.000
	Max. hasítási hely valószínűség	0.000	0.800	0.000	0.000	0.000
	Hasítási hely	-1 and 0	20 and 21	-1 and 0	17 and 18	28 and 29
	Előrejelzés	Nincs szignálpeptid	Szignálpeptid	Nincs szignálpeptid	Nincs szignálpeptid	Nincs szignálpeptid
Heterogán NN	max. C					
	max. Y					
	max. S					
	mean S					
	D					
Pfam						
		Glikozil hidroláz 81	Glikozil hidroláz 81	Glycosyl hydrolase family 84	Glikozil hidroláz 81	Glikozil hidroláz 81

Faj		<i>Paracoccidioides brasiliensis</i>			<i>Magnaporthe grisea</i>	
Törzs		Pb01	Pb03	Pb18	70-15	70-15
Kód		XP_002797710.1	EEH16341.1	EEH43059.1	XP_368243.2	XP_360114.2
Fehérje hossza (aa)		782	845	592	820	811
reciprok BLASTp első találat		Eng2	Eng1	Eng2	Eng2	Eng2
NN	Szignálpeptid?	NEM	IGEN	IGEN	NEM	IGEN
HMM	Szignálpeptid valószínűsége	0.006	0.443	0.442	0.001	1.000
	Szignál horgony valószínűsége	0.001	0.000	0.000	0.000	0.000
	Max. hasítási hely valószínűség	0.003	0.286	0.286	0.000	0.550
	Hasítási hely	41 and 42	14 and 15	14 and 15	27 and 28	20 and 21
	Előrejelzés	Nincs szignálpeptid	Nincs szignálpeptid	Nincs szignálpeptid	Nincs szignálpeptid	Szignálpeptid
Heterogán NN	max. C					
	max. Y					
	max. S					
	mean S					
	D					
Pfam						
		Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81

Faj		<i>Candida albicans</i>			<i>Saccharomyces cerevisiae</i>			
Törzs		WO-1	SC5314	SC5314	EC1118	RM11-1a	RM11-1a	EC1118
Kód		EEQ42800.1	XP_721293.1	XP_716726.1	CAY82261.1	EDV12201.1	EDV09439.1	CAY81379.1
Fehérje hossza (aa)		1143	1146	734	1117	1117	779	779
reciprok BLASTp első találat		Eng1	Eng1	Eng2	Eng2	Eng2	Eng2	Eng2
NN	Szignálpeptid?	IGEN	IGEN	NEM	IGEN	IGEN	NEM	NEM
HMM	Szignálpeptid valószínűsége	0.998	0.998	0.000	0.998	0.998	0.000	0.000
	Szignál horgony valószínűsége	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Max. hasítási hely valószínűség	0.849	0.849	0.000	0.994	0.994	0.000	0.000
	Hasítási hely	18 and 19	18 and 19	-1 and 0	16 and 17	16 and 17	-1 and 0	-1 and 0
	Előrejelzés	Signal peptide	Szignálpeptid	Nincs szignálpeptid	Szignálpeptid	Szignálpeptid	Nincs szignálpeptid	Nincs szignálpeptid
Heterogán NN	max. C							
	max. Y							
	max. S							
	mean S							
	D							
Pfam		Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81

Faj		<i>Histoplasma capsulatum</i>				<i>Blastomyces dermatitidis</i>
Törzs		G186AR	H88	NAm1	H143	ATCC 18188
Kód		EEH05945.1	EGC49676.1	XP_001538425.1	EER41848.1	EGE85606.1
Fehérje hossza (aa)		848	848	839	797	932
reciprok BLASTp első találat		Eng2	Eng2	Eng2	Eng2	Eng2
NN	Szignálpeptid?	IGEN	IGEN	heterogén	IGEN	IGEN
HMM	Szignálpeptid valószínűsége	0.997	0.997	0.379	0.997	1.000
	Szignál horgony valószínűsége	0.000	0.000	0.002	0.000	0.000
	Max. hasítási hely valószínűség	0.961	0.961	0.323	0.961	0.995
	Hasítási hely	23 and 24	23 and 24	24 and 25	23 and 24	25 and 26
	Előrejelzés	Szignálpeptid	Szignálpeptid	Nincs szignálpeptid	Szignálpeptid	Szignálpeptid
Heterogán NN	max. C			Igen		
	max. Y			Nem		
	max. S			Nem		
	mean S			Nem		
	D			Nem		
Pfam		Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81	Glikozil hidroláz 81

Köszönetnyilvánítás

Szeretnék köszönetet mondani témavezetőmnek, **Sveiczer Ákosnak** rengeteg szakmai és egyéb segítségével, folyamatos támogatásáért, és belém fektetett idejéért és bizalmáért.

Köszönöm másik témavezetőmnek, **Sipiczki Mátyásnak** szakmai segítségét.

Köszönöm **Horváth Annának** a folyamatos együtt működését és segítségét.