

**Doktori (PhD) értekezés tézisei**

**Fungális iker-intronok (stwintronok) evolúciós és  
funkcionális vizsgálata**

Kavalecz Napsugár

Témavezető: Dr. Fekete Erzsébet



**DEBRECENI EGYETEM**

Juhász-Nagy Pál Doktori Iskola

Debrecen, 2020

# 1. Bevezetés

A XX. század második felétől kezdődően rohamosan nőtt a molekuláris biológia és genetika területén szerzett ismereteink száma, melyek napjainkig intenzíven gyarapodnak. Az elmúlt pár tíz évben a kutatók olyan módszereket és eszközöket fejlesztettek ki, melyek lehetővé teszik, hogy célzottan módosítsuk egy-egy sejt vagy organizmus genetikai állományát, illetve válaszokat találjunk bizonyos genetikai betegségek kialakulásának mechanizmusaira.

Az intronok, sőt az intronok felfedezését megelőzően is már számos fontos megállapítás került leírásra a genetikai örökítőanyaggal kapcsolatosan, mely információk elengedhetetlenek voltak ahhoz, hogy a kutatók továbbhaladhassanak a molekuláris biológia, már részben kitaposott ösvényein. Az egyik legjelentősebb felfedezés valószínűleg a DNS szerkezetének meghatározása volt (Watson és Crick, 1953), melyen hosszú éveken át dolgozott James Watson, Francis Crick, Maurice Wilkins és Rosalind Franklin. A négy kutató közül Watson, Crick és Wilkins megosztott ovisi Nobel-díjat vehettek át 1962-ben. Rosalind Franklin munkásságát a tudományos társaság nem honorálta Nobel-díjjal, feltehetőleg részben korai halálának is köszönhetően; annak ellenére, hogy ő készítette el azt a világhírűvé vált röntgendiffrakciós képet, mely alapján a kutatók kikövetkeztethették a DNS kettős hélixének pontos szerkezetét (Maddox, 2013).

1993-ban Richard J. Roberts és Phillip A. Sharp felfedezéseit szintén Nobel-díjjal ismerték el, akik 1977-ben - egymástól függetlenül - mutatták ki adenovírusokkal végzett kísérleteik során, hogy a gének több, egymástól nem kódoló szakasszokkal elválasztott szegmensből épülnek fel (Berget és mtsai., 1977; Chow és mtsai., 1977). A kutatók azt is megállapították, hogy ezek a DNS darabok nem képezik az érett mRNS részét, valamilyen módon abból eltávolításra kerülnek.

Mivel mindeközéig a kutatók úgy vélték, az mRNS molekulák minden esetben hű másolatai a kódoló DNS szálaknak, az 1977-ben leírt különös jelenség elkezdte foglalkoztatni a tudományos közéletet. Egy évvel később, 1978-ban Walter Gilbert fel is teszi a kérdést: „Why genes in pieces?”; vajon miért állnak darabokból a gének? Gilbert ezeket a kódoló régiókat megszakító szakaszokat „intron”-oknak nevezi, az intragenic region kifejezésből eredeztetve, míg a kódoló szegmenseket „exon”-oknak kereszteli, mint expressed sequences, azaz kifejeződő, átíródó szekvenciák (Gilbert, 1978).

Az intronok különböző típusait és lehetséges funkcióit taglaló publikációk mellett összetett, több intronból felépülő struktúrák leírására is sor került a XX. század végén. 1991-ben Donald W. Copertino és Richard B. Hallick *Euglena gracilis* alga kloroplasztisának genomjában írtak le különleges, két egymásba ékelődött intronból álló szekvenciákat, melyeket „twintron”-oknak neveztek el, a twin (iker) és intron szavak összetételével (Copertino és Hallick, 1991).

A Debreceni Egyetem Biomérnöki Tanszékén jómagam is hasonló struktúrákat vizsgáltam PhD hallgatóként. Mivel az általunk leírt twintronok spliceoszómális intronokból épülnek fel „stwintron”-nak kereszteltük őket, mint spliceoszómális iker-intronok (Flipphi és mtsai., 2013). Ezek az összetett szerkezetek (hasonlóan az alga kloroplasztban leírtakhoz) két, egymást követő kivágási (splicing) lépés során kerülhetnek csak teljes eltávolításra az elsődleges mRNS átiratból.

Az elmúlt néhány évben több, különböző típusú stwintront is sikerült kimutatnunk fonalas gombák egyes génjeiben, beleértve a jelen értekezés alapját képező eredményeket is (Flipphi és mtsai., 2013; Ág és mtsai., 2015; Fekete és mtsai., 2017; Flipphi és mtsai., 2017; Kavalecz és mtsai., 2019). Kezdetben célunk minél több ilyen struktúra leírása volt, majd esetleges funkcióiknak a feltárása. Emellett remek modell rendszereknek tűnnek, melyek segítségével feltérképezhetőek az intronok evolúciós eseményei, valamint egyes alternatív splicing jelenségek lejátszódásához fűződő szerepük.

Úgy vélem, az stwintronok jelenléte nem korlátozódik a gombák országára; nagy valószínűséggel megtalálhatóak magasabb rendű eukarióta szervezetek génjeiben is. Ezen egyszerű organizmusok vizsgálata során elért eredményeink a későbbiekben azonban felhasználhatóak lehetnek akár humán genetikai betegségek, vagy aberráns splicing folyamatok mögött álló jelenségek megértéséhez is.

## 2. Célkitűzések

Az legelső spliceoszómális iker-intron struktúrák felfedezését és az stwintron fogalom bevezetését követően célul tűztük ki minél több összetett megszakító szekvencia leírását. Ennek érdekében megpróbáltunk létrehozni egy stwintron modellt, melynek segítségével célzottan vizsgálhatunk át elérhető gomba genomszekvenciákat, egy adott típusú stwintron után kutatva.

A szinte napról napra gyarapodó szekvenciaadatok és elérhető, megszekvenált genomok pedig lehetővé teszik, hogy megvizsgáljunk ortológ géneket több száz, vagy akár több ezer rokonfajban. Ezt kihasználva célunk volt feltárni az stwintronok és ezáltal a spliceoszómális intronok evolúciójának, eltűnésének, megjelenésének rejtélyeit gombagenomokban.

Stwintronok kimutatása és splicing intermedierjeik azonosítása után szeretnénk volna választ kapni arra a kérdésre is, hogy vajon milyen funkciót töltenek be ezek a különleges, összetett, intronikus struktúrák a genomban. Bírhatnak-e valamilyen szabályozó funkcióval? Esetleg szerepet játszhatnak-e alternatív splicinggal kapcsolatos mechanizmusok lejátszódásában?

### 3. Anyagok és módszerek

#### 3.1. Bioinformatikai módszerek

##### 3.1.1. Putatív stwintron struktúrák keresése teljes genomszekvenciákban

A fungális intronokra jellemző sajátosságokat figyelembe véve, kutatócsoportunk megalkotott egy stwintron modellt, melynek segítségével teljes gombagenomokban kutathatunk putatív iker-intron struktúrák után. A modell alapját *A. nidulans* esetén megállapított konzervált splicing motívumok szekvenciái képezik (Kupfer és mtsai., 2004). A korrigált, Kupfer-féle intron séma segítségével manuálisan létrehoztuk a fungális stwintron modellt, ahol két intron egymásba ágyazódását modellezve öt jellemző szekvenciamotívumot kaptunk. Ezek közül kettő „hibrid” elem, ugyanis egy, vagy több nukleotidot mind a külső, mind a belső intron szekvenciájából magában foglal.

A konzervált splicing motívumokon túl meghatároztuk az egyes elemek közti távolságtartományokat is az alábbi információk birtokában: I. a legrövidebb, *A. nidulans*-ban kimutatott intron 42 nukleotid hosszúságú; II. a minimális távolság az elágazási pont szekvencia és az akceptorhely között 4 nukleotid lehet; III. a donorhely és az elágazási pont szekvencia közti távolság mindig nagyobb, mint az elágazási pont és az akceptorhely közötti; ennek az értéknek 25 és 120 nukleotid közötti tartományt adtunk meg. Az átlagos intronhossz *A. nidulans* genomjában 73 nukleotid, míg a 160 nukleotidnál hosszabb kanonikus intronok igencsak ritkák (Kupfer és mtsai., 2004). Így [D1,2] típusú stwintronok kereséséhez a következő modellt használtuk: GGTRWGYN(25,120)DYTRAYN(4,24)HAGTRWGYN(25,120)DYTRAYN(4,24)HAG teljes genomszekvenciák átvizsgálásához, a Fuzznuc programot alkalmazva (Rice és mtsai., 2000). Egyszerű átalakításokkal (természetesen) más típusú stwintronokra is megalkotható a modell, amit munkánk során ki is használtunk.

##### 3.1.2. Ortológ gének vizsgálata (genome mining) és filogenetikai analízis

A National Centre of Biotechnology Information (NCBI) és a US Department of Energy Joint Genome Institute (JGI) adatbázisokban nyilvánosan elérhető gomba genomszekvenciákat TBLASTN (Altschul és mtsai., 1997) segítségével vizsgáltuk át, bizonyítottan stwintront tartalmazó gének ortológjai után kutatva.

Filogenetikai analízisek elvégzéséhez a fehérje szekvenciákat a MAFFT (Multiple Alignment using Fast Fourier Transform, version7) (Kato és Standley, 2013) programmal hasonlítottuk össze. A fa ághosszainak kiszámításához az Approximate Likelihood Ratio Test-et (Anisimova és Gascuel, 2006) és a PhyML programot használtuk, Chi-2 alapú, parametrikus feltevésrel. A fák megrajzolásában a FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>, version 1.4.3), utólagos szerkesztésében pedig az Adobe Illustrator program volt segítségünkre.

### 3.1.3. Másodlagos fehérjestruktúra meghatározása

A reticulon-like proteint (RtnA) kódoló gének vizsgálatokor a fehérjék másodlagos, transzmembrán hélix struktúráját az online elérhető TMHMM (version 2) szerver (Krogh és mtsai., 2001) segítségével prediktáltuk. A strukturális eltérések kimutatásához a DISOPRED3 (Jones és Cozzetto, 2015), továbbá a „coiled-coil” struktúrák megtalálásához a PCOILS (Gruber és mtsai., 2006) programokat használtuk.

## 3.2. Gombatörzsek

A kísérletes munka során a következő gombatörzseket vizsgáltuk: *Aspergillus nidulans* R21 (ATCC 48756) (Wortman és mtsai., 2009), *Aspergillus niger* ATCC 1015 (Andersen és mtsai., 2011), *Trichoderma reesei* QM 9414 (ATCC 26921) (Vitikainen és mtsai., 2010), *Helminthosporium solani* B-AC-16A (Mattupalli és mtsai., 2014), *Botrytis cinerea* B05.10 (Amsellem és mtsai., 2011), *Neurospora crassa* OR74A (Galagan és mtsai., 2003), *Malbranchea cinnamomea* CBS 343.55 (Morgenstern és mtsai., 2012).

Az egyes gombatörzsek fenntartásához, illetve növesztéséhez használt táptalajok pontos összetétele máshol került leírásra (Pontecorvo és mtsai., 1953; de Vries és mtsai., 2004; Mandels és Andreotti, 1978; Mimee és mtsai., 2011; Fekete és mtsai., 2012; Vogel, 1956; Morgenstern és mtsai., 2012).

### **3.3. Nukleinsav izolálás**

A vizsgált gombatörzseket nukleinsav izoláláshoz 500 ml-es Erlenmeyer-lombikokban növesztettük ki, azokat rázógépben (Infors HT Multitron) 200 rpm-en és az adott törzsre nézve optimális növekedési hőmérsékleten inkubáltuk. A lombikok egyenként 100 ml tápoldatot tartalmaztak, melyeket az adott gombatörzs vegetatív spórájával oltottunk le. A törzs növekedésének gyorsaságától függően, 16-24 órás inkubációt követően „gyűjtöttük be” a biomasszát. A lombikok tartalmát steril Miracloth szűrőtextilen (Calbiochem) keresztül leszűrtük. Az összegyűlt gombamicéliumot jéghideg, steril, ioncserélt vízzel mostuk, majd két papírtörülő réteg között víztelenítettük. A mintákat folyékony nitrogénben azonnal lefagyasztottuk, majd felhasználásig -80°C-os mélyfagyasztóban tároltuk.

A minták porítását dörzsmozsarakban végeztük, folyékony nitrogén folyamatos adagolása mellett. A nukleinsav kivonását Macherey-Nagel NucleoSpin kit-ek (NucleoSpin Plant II-t DNS izoláláshoz és NucleoSpin RNA Plant-et RNS izoláláshoz) segítségével, a gyártó utasításait követve végeztük. A DNS és RNS minták pontos koncentrációját NanoDrop spektrofotométerrel határoztuk meg az izolálást követően.

### **3.4. Reverz transzkripció PCR (RT-PCR) reakciók**

Az RT-PCR reakciók kivitelezéséhez első lépésként cDNS-t szintetizáltunk reverz transzkripció révén. A szintézishez minden esetben pontosan 1 µg RNS-t, illetve oligo(dT) primert és a First Strand cDNA Synthesis Kit-et (Thermo Scientific) használtuk. A PCR reakciók kivitelezéséhez 4 µL cDNS-t alkalmaztunk templátként (25 µL reakcióvégtérfogathoz), melyekről különböző génspecifikus primerek segítségével DreamTaq DNA Polymerase (Thermo Scientific) alkalmazásával szaporítottuk fel a számunkra releváns fragmenseket. Az alkalmazott PCR programok egyes lépései a következők voltak: elődenaturáció (95°C, 2 min), denaturáció (95°C, 30 sec), anelláció (56/60°C, 1 min), elongáció (72°C, 1 min), végső polimerizáció (72°C, 5 min); a denaturációs, anellációs és elongációs lépést 40x ismételve. Az amplifikált PCR termékeket agaróz gélben választottuk el egymástól gélelektroforézist alkalmazva.

Annak érdekében, hogy a prediktált splicing intermediereket ki tudjuk mutatni, olyan primerpárokat alkalmaztunk az RT-PCR reakciók során, melyek nem az érett mRNS templátokról szaporítanak fel terméket. Ezen primerek egyikét jellemzően úgy terveztük meg, hogy az intron-exon csatlakozásnál tapadjanak, néhány nukleotidjukkal a feltételezett stwintron külső intronjának 3' végéhez kötődve. Ezekkel az oligonukleotidokkal általában két termék amplifikálódik a PCR reakció során, melyek közül a kisebb méretű a számunkra releváns. Ez ugyanis azon intermedier jelenlétére utal, mely esetén a belső intron kivágásra került, de a második splicing reakció még nem játszódott le, tehát az stwintron külső intronját még tartalmazza a szekvencia. Ezzel szemben a nagyobb PCR termék az elsődleges mRNS átíratról felszaporodó amplikon. A későbbiekben mindkét fragmens megszekvenáltatása megtörtént. Minden RT-PCR reakciót kétszer végeztünk, egymástól független rázatott lombikos tenyészetekből nyert biomasszából izolált RNS-t használva.

### **3.5. Szekvenálás**

A PCR termékek agaróz gélben történő elválasztása után már prediktálható, hogy mely esetekben szaporodott fel a keresett fragmens, tehát mely struktúrák esetén beszélhetünk valóban stwintronról. Ahhoz, hogy ezt bizonyítsuk, a méretük alapján megfelelőnek ítélt amplikonokat kitisztítottuk, NucleoSpin Gel & PCR Clean-up (Macherey-Nagel) kit segítségével, majd vektorba klónoztuk pGEM-T Easy Vector System I (Promega) rendszert használva. A plazmid izolálást NucleoSpin Plasmid EasyPure kit (Macherey-Nagel) segítségével végeztük, követve a gyártó utasításait.

Minden esetben három, független plazmid klónt küldtünk el szekvenáltatásra (Eurofins Genomics, Ebersberg, Germany). A cég olyan univerzális primerek (pUC/M13 Primer, Forward and Reverse) segítségével végzi a szekvenálást, melyek a vektor konstrukció meghatározott szakaszaival hibridizálnak. Jelen értekezésben bemutatott kísérletek szekvenálási eredményei mentésre kerültek a GenBank-ban, az alábbi sorszámok alatt: KY315812-KY315816; MF612150-MF612153; MK410458-MK410473; MK421638-MK421641.



## 4. Új tudományos eredmények

- I. Megalkottunk egy stwintron modellt, melynek segítségével újabb putatív stwintron struktúrák után kutathatunk, akár teljes genomszekvenciákban.
- II. Bizonyítottuk egy [D1,2] stwintron jelenlétét *A. nidulans dhaoS* génjében, valamint egy alternatív kivágódásra képes, [D1,2]/[A2,3] típusú stwintron létét *T. reesei* ortológ génjében, azonos pozícióban. A felfedezett alternatív stwintron, valamint splicing intermedierjeinek segítségével egy lehetséges magyarázatot adtunk az intronok egy nukleotidos elcsúszásának mechanizmusára, eukarióta genomokban.
- III. *A. nidulans* és *A. niger* egy putatív lipázt kódoló génjében (*lipS*) kimutattunk egy újabb [D1,2] stwintront; *T. reesei* bifunkcionális, a biotin bioszintézishez szükséges enzimet kódoló génjében (*bioDA*) pedig egy [D2,3] stwintron jelenlétét bizonyítottuk. Az stwintront tartalmazó génekről képződő fehérjetermékek molekuláris filogenetikájának segítségével lehetőségünk nyílt vizsgálni a spliceoszómális iker-intronok megjelenését, valamint eltűnését gombagenomokból. Emellett bizonyítottuk, hogy az stwintronok remek modell rendszerek lehetnek a spliceoszómális intron evolúció vizsgálatát illetően.
- IV. *A. nidulans*, egy reticulon-like proteint kódoló génjében (*rtnA*), egy [D5,6] típusú stwintront sikerült kimutatnunk. Bizonyítottuk, hogy az iker-intron struktúra szerepet játszhat alternatív splicing mechanizmusok (exon skipping) lejátszódásában, *A. niger* és *N. crassa* fonalas gombákban.
- V. Az *Onygenales* rend egyes gombafajainak reticulon-like proteint kódoló génjében egy új intron alakult ki, feltehetően intronizáció révén, a [D5,6] stwintron és egy másik, kanonikus intron között, illetve körül. Ez az új struktúra rendkívül komplex, teljes eltávolítása az mRNS-ből négy, egymást követő splicing reakció lejátszódását igényli. Eredményeink bizonyítják, hogy az intronok spliceoszóma általi eltávolítása nem feltétlenül ko-transzkripcionálisan történik.

## 5. Összefoglalás

Az stwintronok olyan speciális megszakító szekvenciák az eukarióta genomban, melyek felépítésüket tekintve két, egymásba ágyazódott spliceoszómális intronból állnak. Az egyik intront (belső), egy másik (külső) intron valamely, a splicing folyamat végbemeneteléhez szükséges szekvenciaelemébe (5' hasítóhely, vagy donor szekvencia / 3' hasítóhely, vagy akceptor szekvencia / lariat, vagy elágazási pont szekvencia) beékelődve találjuk. Mivel a belső intron a külső intron egy, a kivágódás végbemenetele szempontjából esszenciális motívumát szakítja meg, a belső intron eltávolításának mindenképp meg kell előznie a külső intron splicingját. Az stwintronok felfedezését és egyes típusaik karakterizálását követően szeretnénk volna mélyrehatóbban vizsgálni a struktúrák splicing mechanizmusait, funkcióit és evolúcióját.

*A. nidulans dhaoS* génjében bizonyítottuk egy [D1,2] típusú stwintron jelenlétét. Ezen struktúra esetén a belső intron a külső intron donor elemét szakítja meg, annak első és második nukleotidja között. *T. reesei* ortológ génjében, azonos pozícióban pedig egy alternatív kivágódásra képes [D1,2]/[A2,3] stwintront karakterizáltunk, mely esetében két különböző, egymással egyetlen nukleotidban (G) átfedő belső intront fedezhetünk fel. Kísérletesen bizonyítottuk a struktúra [D1,2] és [A2,3] útvonalon történő eltávolítását is, illetve, hogy mindkét út azonos mRNS termékhez vezet. Bizonyítottuk továbbá, hogy az alternatív stwintron egyik, vagy másik belső intronjának alternatív elvesztése közeli, de nem azonos intron pozíciókat eredményezhet a genomban; ezzel egy lehetséges magyarázattal szolgálva az egy nukleotiddal történő intron elcsúszás jelenségére.

Az egyre nagyobb számban rendelkezésre álló genomszekvencia adatoknak hála, a *lipS* és *bioDA* génekben leírt [D1,2], valamint [D2,3] típusú stwintronokat, azok kialakulását, vagy éppen eltűnését vizsgálni tudtuk gombák egész rendjein és osztályain belül. A *lipS* ortológok vizsgálata során azt tapasztaltuk, hogy az stwintron pozíciót egyes fajok esetén egy normál intron foglalja el. Ebből arra következtettünk, hogy a [D1,2] stwintron egy új intron, már korábban jelenlévő intronba történő beágyazódásával jöhetett létre. *Leotiomycetes*-ek és *Eurotiomycetes*-ek között viszont az stwintron teljes eltűnésére is találunk példát.

Ezzel szemben a *bioDA* ortológok vizsgálatánál azt láttuk, hogy az stwintron pozíciót bizonyos esetekben szintén egy normál intron foglalja el, viszont intron veszteségi események lejátszódásával, amit egyes *Sordariomycetes* és *Botryosphaerales* fajoknál mutattunk ki, a [D2,3] stwintron csak a belső intronját veszítette el, egy normál U2 intront hátrahagyva.

Több fonalas gombafaj ortológ génjében is kísérletesen igazoltuk, az először *A. nidulans* *rtnA* génjében kimutatott [D5,6] típusú stwintron jelenlétét. *A. niger* és *N. crassa* esetén bizonyítottuk, hogy a struktúra egy alternatív splicing esemény lejátszódásában is részt vehet, mégpedig a szekvenciában mellette elhelyezkedő (downstream irányban) exon kivágása révén. Amennyiben az stwintron belső intronja alternatív 3' hasítóhely használatával kerül eltávolításra, a külső intron, a szekvenciában utána következő normál intron akceptorhelyét veszi igénybe, ezzel egyetlen nagy intronként kivágódva, magában foglalva az említett exonikus szakaszt is.

Az *rtnA* ortológok további vizsgálata során kimutattuk, hogy 29 *Onygenales* faj esetén egy többlépéses splicing mechanizmus játszódik le a transzkriptum érése során; ugyanis a gén egy komplex megszakító szekvenciát tartalmaz, mely magában foglalja a [D5,6] stwintront, a szomszédos kanonikus U2 intront, és egy ezeket körülölelő másodrendű intront is. A teljes struktúra eltávolításához szükséges négy splicing reakció kivágódási közteseit kimutattuk *M. cinnamomea* fonalas gombában.

Végezetül elmondható, hogy a megalkotott stwintron modell segítségével, fonalas gombák különböző génjeiben *in silico* újabb spliceoszómális iker-intronokat sikerült dekteltátunk és splicing közteseik kimutatása révén létezésüket bizonyítanunk. Emellett bebizonyosodott, hogy ezek a különleges stwintron struktúrák remek modell rendszerek lehetnek a spliceoszómális intronok eredetét, evolúcióját, genomon belüli mobilitását illetően. Részben választ kaptunk arra a kérdésre is, hogy vajon mi lehet a feladatuk, funkciójuk a genomban, hiszen kimutattuk, hogy szerepet játszhatnak alternatív splicing mechanizmusok (jelen esetben exon skipping) lejátszódásában. Biztos vagyok benne, hogy emellett még számos más funkcióval, akár a génexpresszióval kapcsolatos regulációs szereppel bírhatnak, de ezeknek a rejtélyeknek a megoldását már másra hagyom.

## 6. Irodalomjegyzék

- Ág, N., Flippi, M., Karaffa, L., Scazzocchio, C., & Fekete, E. (2015). Alternatively spliced, spliceosomal twin introns in *Helminthosporium solani*. *Fungal Genetics and Biology*, 85, 7-13.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), 3389-3402.
- Amselem, J., Cuomo, C. A., Van Kan, J. A., Viaud, M., Benito, E. P., Couloux, A., ... & Fournier, E. (2011). Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS genetics*, 7(8), e1002230; 10.1371/journal.pgen.1002230
- Andersen, M. R., Salazar, M. P., Schaap, P. J., Van De Vondervoort, P. J., Culley, D., Thykaer, J., ... & Berka, R. M. (2011). Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. *Genome research*, 21(6), 885-897.
- Anisimova, M., & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic biology*, 55(4), 539-552.
- Berget, S. M., Moore, C., & Sharp, P. A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences*, 74(8), 3171-3175.
- Chow, L. T., Gelinas, R. E., Broker, T. R., & Roberts, R. J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, 12(1), 1-8.
- Copertino, D. W., & Hallick, R. B. (1991). Group II twintron: an intron within an intron in a chloroplast cytochrome b-559 gene. *The EMBO journal*, 10(2), 433-442.
- de Vries, R. P., Burgers, K., van de Vondervoort, P. J., Frisvad, J. C., Samson, R. A., & Visser, J. (2004). A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. *Applied and Environmental Microbiology*, 70(7), 3954-3959.
- Fekete, É., Fekete, E., Irinyi, L., Karaffa, L., Árnayasi, M., Asadollahi, M., & Sándor, E. (2012). Genetic diversity of a *Botrytis cinerea* cryptic species complex in Hungary. *Microbiological research*, 167(5), 283-291.
- Flippi, M., Fekete, E., Ág, N., Scazzocchio, C., & Karaffa, L. (2013). Spliceosome twin introns in fungal nuclear transcripts. *Fungal genetics and biology*, 57, 48-57.
- Gilbert, W. (1978). Why genes in pieces? *Nature*, 271(5645), 501-501.
- Gruber, M., Söding, J., & Lupas, A. N. (2006). Comparative analysis of coiled-coil prediction methods. *Journal of structural biology*, 155(2), 140-145.
- Jones, D. T., & Cozzetto, D. (2015). DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*, 31(6), 857-863.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4), 772-780.

- Krogh, A., Larsson, B., Von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology*, 305(3), 567-580.
- Kupfer, D. M., Drabenstot, S. D., Buchanan, K. L., Lai, H., Zhu, H., Dyer, D. W., Roe, B. A. & Murphy, J. W. (2004). Introns and splicing elements of five diverse fungi. *Eukaryotic cell*, 3(5), 1088-1100.
- Maddox, B. (2003). The double helix and the 'wronged heroine'. *Nature*, 421(6921), 407-408.
- Mandels, M. M. & Andreotti, R. E. (1978) Problems and challenges in the cellulose to cellulase fermentation. *Process Biochemistry*, 13, 6–13.
- Mattupalli, C., Glasner, J. D., & Charkowski, A. O. (2014). A draft genome sequence reveals the *Helminthosporium solani* arsenal for cell wall degradation. *American journal of potato research*, 91(5), 517-524.
- Mimee, B., Avis, T. J., Boivin, S., Jabaji, S., & Tweddell, R. J. (2011). Effect of iron and nitrogen on the development of *Helminthosporium solani* and potato silver scurf. *Canadian journal of plant pathology*, 33(4), 506-511.
- Morgenstern, I., Powlowski, J., Ishmael, N., Darmond, C., Marqueteau, S., Moisan, M. C., Quenneville, G. & Tsang, A. (2012). A molecular phylogeny of thermophilic fungi. *Fungal Biology*, 116(4), 489-502.
- Pontecorvo, G., Roper, J. A., Chemmons, L. M., MacDonald, K. D., & Bufton, A. W. J. (1953). The genetics of *Aspergillus nidulans*. *Advances in genetics* 5, 141-238.
- Rice, P., Longden, I., & Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends in Genetics*, 16, 276-277.
- Vitikainen, M., Arvas, M., Pakula, T., Oja, M., Penttilä, M., & Saloheimo, M. (2010). Array comparative genomic hybridization analysis of *Trichoderma reesei* strains with enhanced cellulase production properties. *BMC genomics*, 11(1), 441, 10.1186/1471-2164-11-441
- Vogel, H. J. (1956). A convenient growth medium for *Neurospora crassa*. *Microbial Genet. Bull.*, 13, 42-43
- Watson, J. D., & Crick, F. H. (1953). Molecular structure of nucleic acids. *Nature*, 171(4356), 737-738.
- Wortman, J. R., Gilsenan, J. M., Joardar, V., Deegan, J., Clutterbuck, J., Andersen, M. R., ... & Von Döhren, H. (2009). The 2008 update of the *Aspergillus nidulans* genome annotation: a community effort. *Fungal Genetics and Biology*, 46(1), S2-S13



Nyilvántartási szám: DEENK/253/2020.PL  
Tárgy: PhD Publikációs Lista

Jelölt: Kavalecz Napsugár  
Doktori Iskola: Juhász-Nagy Pál Doktori Iskola  
MTMT azonosító: 10060293

### **A PhD értekezés alapjául szolgáló közlemények**

#### Idegen nyelvű tudományos közlemények külföldi folyóiratban (3)

1. **Kavalecz, N.**, Ág, N., Karaffa, L., Scazzocchio, C., Flippi, M., Fekete, E.: A spliceosomal twin intron (stwintron) participates in both exon skipping and evolutionary exon loss.  
*Sci. Rep.* 9 (1), 1-11, 2019. EISSN: 2045-2322.  
DOI: <http://dx.doi.org/10.1038/s41598-019-46435-x>  
IF: 3.998
2. Fekete, E., Flippi, M., Ág, N., **Kavalecz, N.**, Cerqueira, G. C., Scazzocchio, C., Karaffa, L.: A mechanism for a single nucleotide intron shift.  
*Nucleic Acids Res.* 45 (15), 9085-9092, 2017. ISSN: 0305-1048.  
DOI: <http://dx.doi.org/10.1093/nar/gkx520>  
IF: 11.561
3. Flippi, M., Ág, N., Karaffa, L., **Kavalecz, N.**, Cerqueira, G. C., Scazzocchio, C., Fekete, E.:  
Emergence and loss of spliceosomal twin introns.  
*Fungal Biol Biotechnol.* 4 (1), 7, 2017. EISSN: 2054-3085.  
DOI: <http://dx.doi.org/10.1186/s40694-017-0037-y>





## További közlemények

### Idegen nyelvű tudományos közlemények külföldi folyóiratban (2)

4. Ág, N., **Kavalecz, N.**, Péntes, F., Karaffa, L., Scazzocchio, C., Flippi, M., Fekete, E.: Complex intron generation in the yeast genus *Lipomyces*.  
*Sci. Rep. 10* (6022), 1-10, 2020. EISSN: 2045-2322.  
DOI: <http://dx.doi.org/10.1038/s41598-020-63239-6>  
IF: 3.998 (2019)
5. Fekete, E., Orosz, A., Kulcsár, L., **Kavalecz, N.**, Flippi, M., Karaffa, L.: Characterization of a second physiologically relevant lactose permease gene (*lacpB*) in *Aspergillus nidulans*.  
*Microbiology (Reading, Engl.)*. 162, 837-847, 2016. ISSN: 1350-0872.  
DOI: <http://dx.doi.org/10.1099/mic.0.000267>  
IF: 2.151

**A közlő folyóiratok összesített impakt faktora: 21,708**

**A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre):  
15,559**

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudományometriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.

Debrecen, 2020.08.31.



**Short thesis for the degree of doctor of philosophy (PhD)**

**Investigation of the evolution and function  
of fungal stwintrons**

by Napsugár Kavalecz

Supervisor: Dr. Erzsébet Fekete



UNIVERSITY OF DEBRECEN  
Juhász-Nagy Pál Doctoral School

Debrecen, 2020



# 1. Introduction

From the second half of the 20<sup>th</sup> century, our scientific knowledge in the field of molecular biology and genetics has intensively increased. With the help of recently developed tools and methods we have the opportunity to modify the genetic material of a single cell or an entire organism. Furthermore, we may have the chance to reveal the mechanisms responsible for genetic diseases.

Several crucial statements, connecting to genetic material, had been set up before the discovery of introns and even exons, which information were necessary for scientists to go further on the paths of molecular biology. One of the most remarkable discoveries probably was the determination of the structure of DNA (Watson and Crick, 1953), for what we can be grateful to James Watson, Francis Crick, Maurice Wilkins and Rosalind Franklin. Three of them (namely Watson, Crick and Wilkins) were awarded by Nobel-prize in 1962 for their discovery. Scientific achievements of Rosalind Franklin has never been honoured by a Nobel-prize, in spite of the fact that she created that world-famous X-ray photograph, which represents the structure of the DNA double helix (Maddox, 2013).

In 1993 Richard J. Roberts' and Phillip A. Sharp's achievement was also awarded by Nobel-prize. The two scientists - independently from each other - showed that coding regions of a gene are separated by non-coding sequences, during their work with adenoviruses (Berget et al., 1977; Chow et al., 1977). They also concluded that these segments aren't part of the coding mRNA sequence, they are removed from this during the maturation process.

While, it was believed that mRNA molecules are exact copy of the coding DNA, the revealed phenomenon surprised the scientific community in 1977. One year later, in 1978, Walter Gilbert asks the question: „Why genes in pieces?“ and he designates these mysterious non-coding sequences „introns“ (intragenic regions), while the coding sequences, which participate in the formation of the mature mRNA, he calls „exons“ (expressed regions) (Gilbert et al., 1978).

In the end of the 20<sup>th</sup> century, next to publications talking about possible functions and different groups of introns, we can find some, which reveal complex intron structures consist of two introns embedded each other. In 1991, Donald W. Copertino and Richard B. Hallick showed an extraordinary intron structure, containing two introns localized in the chloroplast

genome of *Euglena gracilis* algae. They named this unique element „twintron” originating from „twin” and „intron” (Copertino and Hallick, 1991).

As a Phd student, I am also investigating structures similar to Copertino and Hallick’s twintron, at the University of Debrecen, Department of Biochemical Engineering. These twintrons consist of spliceosomal introns, so we call them „stwintrons” (spliceosomal twintrons) (Flipphi et al., 2013). These structures require two consecutive splicing reactions for their total removal from the mRNA, just like the one found in the algae chloroplast.

In the last few years, many type of stwintrons were found in different fungal genes, including results described in this thesis (Flipphi et al., 2013; Ág et al., 2015; Fekete et al., 2017; Flipphi et al., 2017; Kavalecz et al., 2019). In the beginning, our aim was to identify as much stwintrons as we can, then we have tried to reveal their possible functions in the eukaryotic genome. We proved that stwintrons can serve as model systems to study spliceosomal intron evolution and they may play important roles connecting to alternative splicing mechanisms.

In my view, we probably could find stwintron structures in the genome of higher eukaryotes not only in filamentous fungi. Albeit, discoveries reached via studying these simple organisms can facilitate the improvement of solutions for the cure of genetic diseases or aberrant splicing mechanisms, also in human.

## 2. Objectives

After the discovery of the first complex intron structures and introducing the term „stwintron” we tried to find and prove as much stwintrons as we can. Our aim was to devise a bioinformatic tool that enable us to identify putative stwintrons in whole genome sequences.

The availability of complete genome sequences provides a unique chance to investigate orthologous genes in hundreds of fungal species. Exploiting this opportunity we decided to study spliceosomal intron evolution, intron gain and loss in conserved genes throughout the whole fungal kingdom.

After detection of stwintrons via their splicing intermediates, we tried to find answers for the following questions: What could be the role of stwintron structures in the eukaryotic genome? Do they have any regulatory function? Do they participate in alternative splicing mechanisms?

## **3. Materials and methods**

### **3.1. Bioinformatic methods**

#### **3.1.1. Tool to search putative stwintrons in whole genome sequences**

We defined five degenerated sequence motifs for the donor-, acceptor- and lariat branch point sequence elements within a stwintron, including the two hybrid motifs characteristic for the stwintron type, consistent of nt of the external as well as of the internal intron. These motifs are based on a statistical consensus for the three conserved elements in *A. nidulans* (Kupfer et al., 2004) although we used a more relaxed consensus.

Furthermore, we defined distance ranges separating these five motifs conforming three principles rooted in our experience in calling intron-exon structure: I. the minimum length of an *A. nidulans* intron is 42 nt; II. the minimum distance between the 6-nt lariat branch point element and the 3-nt acceptor element is 4 nt; and III. the distance between the 6-nt donor element at 5' and the lariat branch point element is always bigger than the distance between the latter and the acceptor at 3'. We set the distance range between the donor- and lariat branch point elements from 25 to 120 nt. The mean intron length is reported to be 73 nt in *A. nidulans* (Kupfer et al., 2004). For a screen for putative [D1,2] stwintrons, these criteria led to the following sequence pattern: GGTRWGYN(25,120)DYTRAYN(4,24)HAGTRWGYN(25,120)DYTRAYN(4,24)HAG which was searched for in whole genome sequences using the Fuzznuc application (Rice et al., 2000). With modification of the model we have the opportunity to search for different type of stwintrons.

#### **3.1.2. Genome mining and phylogenetic analyses**

TBLASTN screens were performed in public DNA databases at the websites of the NCBI (National Centre of Biotechnology Information) and JGI (US Department of Energy Joint Genome Institute) using the online facilities (Altschul et al., 1997).

For phylogenetic analyses, proteins were aligned with Multiple Alignment using Fast Fourier Transform (MAFFT; version 7) (Katoh and Standley, 2013), applying the E-INS-i algorithm and BLOSUM similarity matrixes. Approximate Likelihood Ratio Tests provide

statistical branch support (Anisimova and Gascuel, 2006) and were integrally calculated by PhyML with the Chi2-based parametric provided. The tree was visualized with the FigTree drawing program (<http://tree.bio.ed.ac.uk/software/figtree>) and annotated with Adobe Illustrator.

### **3.1.3. Secondary structure predictions**

Transmembrane helices of the RtnA protein were predicted online using the TMHMM (version 2) server (Krogh et al., 2001). Structural disorder was predicted with DISOPRED3 (Jones and Cozzetto, 2015) and coiled-coil structures with PCOILS (Gruber et al., 2006).

## **3.2. Fungal strains**

During our laboratory experiments we used the following fungal strains: *Aspergillus nidulans* R21 (ATCC 48756) (Wortman et al., 2009), *Aspergillus niger* ATCC 1015 (Andersen et al., 2011), *Trichoderma reesei* QM 9414 (ATCC 26921) (Vitikainen et al., 2010), *Helminthosporium solani* B-AC-16A (Mattupalli et al., 2014), *Botrytis cinerea* B05.10 (Amselem et al., 2011), *Neurospora crassa* OR74A (Galagan et al., 2003), *Malbranchea cinnamomea* CBS 343.55 (Morgenstern et al., 2012).

Their maintenance and standardized growth medium compositions are described elsewhere (Pontecorvo et al., 1953; de Vries et al., 2004; Mandels and Andreotti, 1978; Mimee et al., 2011; Fekete et al., 2012; Vogel, 1956; Morgenstern et al., 2012).

## **3.3. Nucleic acid isolation**

Fungal strains were grown in 500 ml Erlenmeyer-flasks with 100 ml of medium seeded with vegetative spore inoculums, in a rotary shaker (Infors HT Multitron) at 200 rpm for 16-24 hours. Mycelia were harvested by filtration over sterile cheese cloth. The biomass was washed with distilled water, frozen and ground to powder under liquid nitrogen. For the extraction of genomic DNA and total RNA, Macherey Nagel NucleoSpin kits (NucleoSpin Plant II and NucleoSpin RNA Plant, respectively) were used according to the manufacturer's instructions.

### **3.4. Reverse transcription PCR (RT-PCR)**

Reverse transcription was primed off 1 µg of total RNA with Oligo(dT) as a primer using the Transcriptor High Fidelity cDNA Synthesis Kit (Roche). Polymerase chain reaction (PCR) reactions were performed in a 25 µl volume containing 4 µl of single strand cDNA, using gene-specific oligonucleotides as primers and DreamTaq DNA Polymerase (Thermo Scientific). Cycling conditions after initial denaturation at 95°C for 2 min were: 40 cycles of 95°C for 30 sec, 56/60°C for 1 min and 72°C for 1 min, followed by one post-cyclic elongation at 72°C for 5 min. Amplified fragments were resolved in native agarose gels.

To confirm the existence of the predicted stwintron splicing intermediates, we used an extraordinary RT-PCR approach with primer pairs that do not amplify DNA off mature mRNA template. One of the primers eclipses an intron-exon junction of the putative external intron and contains few nt at its 3' end that do not basepair when the external intron is absent. This protocol usually yields two PCR fragments of defined sizes of which the smaller one corresponds to the splicing intermediate. Both fragments were processed and sequenced. All RT-PCR experiments were done in duplicate, starting with biomass from two independent liquid cultures.

### **3.5. cDNA sequencing**

After gelelectrophoresis, cDNA was purified with NucleoSpin Gel & PCR Clean-up (Macherey-Nagel) and subsequently cloned using pGEM-T Easy Vector System I (Promega). Three independent plasmid clones were sequenced (Eurofins Genomics, Ebersberg, Germany) over both strands using universal primers hybridizing to the vector (pUC/M13 Primer, Forward and Reverse). cDNA sequences were deposited at GenBank under the following accession numbers: KY315812-KY315816; MF612150-MF612153; MK410458-MK410473; MK421638-MK421641.

## 4. New scientific results

- I. We devised a bioinformatic tool that enabled us to identify putative stwintron structures in whole genome sequences.
- II. We detected a [D1,2] stwintron in the *dhaoS* gene of *A. nidulans* and an alternatively spliced [D1,2]/[A2,3] stwintron in the orthologous gene of *T. reesei*. We elucidated a novel mechanism of 1-nt intron shift in eukaryotic nuclear transcripts with the alternatively spliced, [D1,2]/[A2,3] stwintron as the key intermediate.
- III. We evidenced a new [D1,2] stwintron in a putative lipase gene (*lipS*) in *A. nidulans* and *A. niger* and a [D2,3] stwintron in the bifunctional biotin biosynthesis gene (*bioDA*) in *T. reesei*. Molecular phylogeny of the peptide products were used to monitor the emergence and disappearance of ancient spliceosomal twin introns. We thus demonstrated that stwintrons can serve as model systems to study spliceosomal intron evolution.
- IV. We detected a [D5,6] stwintron in a reticulon-like protein coding gene (*rtnA*) in *A. nidulans*. We showed that this stwintron is implicated in a mechanism of alternative splicing (exon skipping) by its ability to use an alternative acceptor for its internal intron, as we have showed it in *A. niger* and *N. crassa*.
- V. In the order of *Onygenales*, a new intron has evolved by intronisation around and in between the pre-extant [D5,6] stwintron and the downstream intron in their reticulon gene. This new intervening sequence is really complex, its removal necessitates four consecutive splicing reactions. Our results confirm that U2 introns are not excised co-transcriptionally in any case.

## 5. Summary

Stwintrons are unconventional intervening sequences in the eukaryotic genome. They consist of two spliceosomal introns; one of the introns (internal intron) is nested within one of the three canonical splicing motifs (5' splice site or donor sequence / 3' splice site or acceptor sequence / the lariat sequence around the branch point adenosine) of another intron (external intron). While the internal intron interrupts a sequence essential for splicing in the external intron, its removal has to take place first. After the discovery of stwintron structures our aims were to examine the function, evolution and splicing mechanisms of these unique intervening sequences.

We evidenced the existence of a [D1,2] stwintron in the *dhaoS* gene of *A. nidulans*, in which an internal intron interrupts the donor sequence of an external intron, between its first and second nucleotide. An alternative [D1,2]/[A2,3] stwintron has also been characterized in the orthologous gene of *T. reesei*. In this case we can detect two different internal introns, which overlap in one single nucleotide (G). The two alternative splicing pathways of this [D1,2]/[A2,3] stwintron has been experimentally proved and we concluded that both routes result the same mRNA product. Furthermore, the alternative loss of the internal intron can result close, but not identical intron positions in the genome; revealing a possible explanation for the one-nucleotide intron shift phenomenon.

Thanks to the increasing number of accessible genome sequences, we had the opportunity to study the emergence, loss and evolution of the *lipS* [D1,2] and the *bioDA* [D2,3] stwintrons throughout whole taxa and classes of fungi. We showed that the stwintron position is occupied by a canonical intron in some *lipS* orthologues. Thus, we hypothesized that this [D1,2] stwintron emerged by the appearance of a new intron within another ancient intron. However, we can find examples of total disappearance of the stwintron among *Eurotiomycetes*.

We experienced a different scenario regarding to *bioDA* orthologues: the stwintron position can also be occupied by a normal intron, however intron loss can take place among *Sordariomycetes* and *Botryosphaeraiales* species, which results the presence of a canonical U2 intron in the stwintron position by loss of the internal intron.



We proved the presence of a [D5,6] stwintron (first described in *A. nidulans rtnA* gene) in orthologous gene of many filamentous fungi. We showed the relevance of the stwintron regarding to alternative splicing events, namely exon skipping, in *A. niger* and *N. crassa*. If the internal intron of the stwintron is using an alternative 3' splice site, the external intron has to use the acceptor site of the neighbouring canonical intron in the gene sequence, thus the external intron is cleaved together with this U2 intron and a small exon in between.

In 29 species of *Onygenales* a complex intervening sequence has been revealed including the [D5,6] stwintron, the neighbouring U2 intron and a „second order intron”. The total removal of this complex structure requires four consecutive splicing reactions as it has been experimentally proved in the transcript of the *M. cinnamomea* orthologous gene.

Summarizing the results we can say that new stwintrons were found *in silico* using the recently created model and experimentally confirmed via the detection of splicing intermediates. We proved that stwintrons can serve as excellent model systems regarding to spliceosomal intron loss or emergence, mobility and evolution. Furthermore, we showed that stwintrons can participate in alternative splicing events (in this case exon skipping), so we are closer to solve the mystery about stwintrons and their role in the genome. In my view, these unique structures would have many other functions connected to gene expression and regulation among others. However, I'll let others to solve these mysteries in the future.

## 6. References

- Ág, N., Flipphi, M., Karaffa, L., Scazzocchio, C., & Fekete, E. (2015). Alternatively spliced, spliceosomal twin introns in *Helminthosporium solani*. *Fungal Genetics and Biology*, 85, 7-13.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), 3389-3402.
- Amselem, J., Cuomo, C. A., Van Kan, J. A., Viaud, M., Benito, E. P., Couloux, A., ... & Fournier, E. (2011). Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS genetics*, 7(8), e1002230; 10.1371/journal.pgen.1002230
- Andersen, M. R., Salazar, M. P., Schaap, P. J., Van De Vondervoort, P. J., Culley, D., Thykaer, J., ... & Berka, R. M. (2011). Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. *Genome research*, 21(6), 885-897.
- Anisimova, M., & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic biology*, 55(4), 539-552.
- Berget, S. M., Moore, C., & Sharp, P. A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences*, 74(8), 3171-3175.
- Chow, L. T., Gelinas, R. E., Broker, T. R., & Roberts, R. J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, 12(1), 1-8.
- Copertino, D. W., & Hallick, R. B. (1991). Group II twintron: an intron within an intron in a chloroplast cytochrome b-559 gene. *The EMBO journal*, 10(2), 433-442.
- de Vries, R. P., Burgers, K., van de Vondervoort, P. J., Frisvad, J. C., Samson, R. A., & Visser, J. (2004). A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. *Applied and Environmental Microbiology*, 70(7), 3954-3959.
- Fekete, É., Fekete, E., Irinyi, L., Karaffa, L., Árnayasi, M., Asadollahi, M., & Sándor, E. (2012). Genetic diversity of a *Botrytis cinerea* cryptic species complex in Hungary. *Microbiological research*, 167(5), 283-291.
- Flipphi, M., Fekete, E., Ág, N., Scazzocchio, C., & Karaffa, L. (2013). Spliceosome twin introns in fungal nuclear transcripts. *Fungal genetics and biology*, 57, 48-57.
- Gilbert, W. (1978). Why genes in pieces? *Nature*, 271(5645), 501-501.
- Gruber, M., Söding, J., & Lupas, A. N. (2006). Comparative analysis of coiled-coil prediction methods. *Journal of structural biology*, 155(2), 140-145.
- Jones, D. T., & Cozzetto, D. (2015). DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*, 31(6), 857-863.

- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4), 772-780.
- Krogh, A., Larsson, B., Von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology*, 305(3), 567-580.
- Kupfer, D. M., Drabenstot, S. D., Buchanan, K. L., Lai, H., Zhu, H., Dyer, D. W., Roe, B. A. & Murphy, J. W. (2004). Introns and splicing elements of five diverse fungi. *Eukaryotic cell*, 3(5), 1088-1100.
- Maddox, B. (2003). The double helix and the 'wronged heroine'. *Nature*, 421(6921), 407-408.
- Mandels, M. M. & Andreotti, R. E. (1978) Problems and challenges in the cellulose to cellulase fermentation. *Process Biochemistry*, 13, 6–13.
- Mattupalli, C., Glasner, J. D., & Charkowski, A. O. (2014). A draft genome sequence reveals the *Helminthosporium solani* arsenal for cell wall degradation. *American journal of potato research*, 91(5), 517-524.
- Mimee, B., Avis, T. J., Boivin, S., Jabaji, S., & Tweddell, R. J. (2011). Effect of iron and nitrogen on the development of *Helminthosporium solani* and potato silver scurf. *Canadian journal of plant pathology*, 33(4), 506-511.
- Morgenstern, I., Powlowski, J., Ishmael, N., Darmond, C., Marqueteau, S., Moisan, M. C., Quenneville, G. & Tsang, A. (2012). A molecular phylogeny of thermophilic fungi. *Fungal Biology*, 116(4), 489-502.
- Pontecorvo, G., Roper, J. A., Chemmons, L. M., MacDonald, K. D., & Bufton, A. W. J. (1953). The genetics of *Aspergillus nidulans*. *Advances in genetics* 5, 141-238.
- Rice, P., Longden, I., & Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends in Genetics*, 16, 276-277.
- Vitikainen, M., Arvas, M., Pakula, T., Oja, M., Penttilä, M., & Saloheimo, M. (2010). Array comparative genomic hybridization analysis of *Trichoderma reesei* strains with enhanced cellulase production properties. *BMC genomics*, 11(1), 441, 10.1186/1471-2164-11-441
- Vogel, H. J. (1956). A convenient growth medium for *Neurospora crassa*. *Microbial Genet. Bull.*, 13, 42-43
- Watson, J. D., & Crick, F. H. (1953). Molecular structure of nucleic acids. *Nature*, 171(4356), 737-738.
- Wortman, J. R., Gilsenan, J. M., Joardar, V., Deegan, J., Clutterbuck, J., Andersen, M. R., ... & Von Döhren, H. (2009). The 2008 update of the *Aspergillus nidulans* genome annotation: a community effort. *Fungal Genetics and Biology*, 46(1), S2-S13



Registry number: DEENK/253/2020.PL  
Subject: PhD Publication List

Candidate: Napsugár Kavalecz

Doctoral School: Pál Juhász-Nagy Doctoral School of Biology and Environmental Sciences

MTMT ID: 10060293

### List of publications related to the dissertation

#### Foreign language scientific articles in international journals (3)

1. **Kavalecz, N.**, Ág, N., Karaffa, L., Scazzocchio, C., Flipphi, M., Fekete, E.: A spliceosomal twin intron (stwintron) participates in both exon skipping and evolutionary exon loss.  
*Sci. Rep.* 9 (1), 1-11, 2019. EISSN: 2045-2322.  
DOI: <http://dx.doi.org/10.1038/s41598-019-46435-x>  
IF: 3.998
2. Fekete, E., Flipphi, M., Ág, N., **Kavalecz, N.**, Cerqueira, G. C., Scazzocchio, C., Karaffa, L.: A mechanism for a single nucleotide intron shift.  
*Nucleic Acids Res.* 45 (15), 9085-9092, 2017. ISSN: 0305-1048.  
DOI: <http://dx.doi.org/10.1093/nar/gkx520>  
IF: 11.561
3. Flipphi, M., Ág, N., Karaffa, L., **Kavalecz, N.**, Cerqueira, G. C., Scazzocchio, C., Fekete, E.:  
Emergence and loss of spliceosomal twin introns.  
*Fungal Biol Biotechnol.* 4 (1), 7, 2017. EISSN: 2054-3085.  
DOI: <http://dx.doi.org/10.1186/s40694-017-0037-y>





## List of other publications

### Foreign language scientific articles in international journals (2)

4. Ág, N., **Kavalecz, N.**, Péntzes, F., Karaffa, L., Scazzocchio, C., Flippi, M., Fekete, E.: Complex intron generation in the yeast genus *Lipomyces*.  
*Sci. Rep.* 10 (6022), 1-10, 2020. EISSN: 2045-2322.  
DOI: <http://dx.doi.org/10.1038/s41598-020-63239-6>  
IF: 3.998 (2019)
5. Fekete, E., Orosz, A., Kulcsár, L., **Kavalecz, N.**, Flippi, M., Karaffa, L.: Characterization of a second physiologically relevant lactose permease gene (*lacpB*) in *Aspergillus nidulans*.  
*Microbiology (Reading, Engl.)*. 162, 837-847, 2016. ISSN: 1350-0872.  
DOI: <http://dx.doi.org/10.1099/mic.0.000267>  
IF: 2.151

**Total IF of journals (all publications): 21,708**

**Total IF of journals (publications related to the dissertation): 15,559**

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

31 August, 2020



