

Alternativno izrezivanje introna u biljaka

Tadić, Vanja

Undergraduate thesis / Završni rad

2012

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:108910>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-04-17**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
BIOLOŠKI ODSJEK

ALTERNATIVNO IZREZIVANJE INTRONA U BILJAKA

ALTERNATIVE SPLICING IN HIGHER PLANTS

SEMINARSKI RAD

Vanja Tadić

Preddiplomski studij molekularne biologije

(Undergraduate Study of Molecular Biology)

Mentor: *doc. dr. sc.* Nataša Bauer

Zagreb, 2012.

SADRŽAJ

1. UVOD.....	4
2. ALTERNATIVNO IZREZIVANJE INTRONA U BILJAKA.....	7
3. CILJ RADA.....	12
4. MATERIJALI I METODE.....	13
5. REZULTATI.....	14
6. ZAKLJUČAK.....	24
7. SAŽETAK/ <i>SUMMARY</i>	26
8. LITERATURA.....	28

Korištene kratice:

snRNPs = *small nuclear ribonucleoproteins*

ESTs = *expressed sequence tags*

RT-PCR = *reverse transcriptase-mediated PCR*

NMD = *nonsense mediated decay*

NGS = *next generation sequencing*

RRM = *RNA recognition motif*

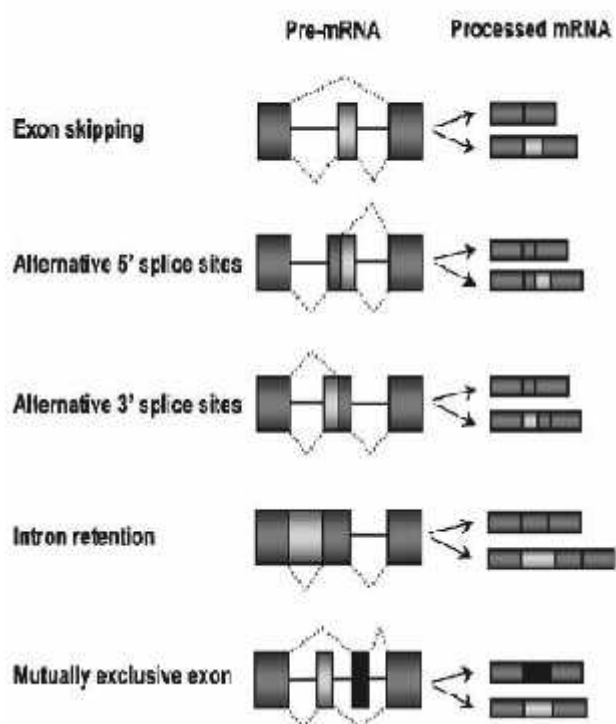
1. UVOD

Prilikom transkripcije DNA u mRNA RNA-polimerazama, transkribiraju se i kodirajuće DNA sekvence (egzoni) i nekodirajuće DNA sekvence (introni). Izrezivanje introna, tzv. *splicing*, dio je posttranskripcijske dorade pre-mRNA koja uključuje dodavanje kape na 5' kraju i poliadenilaciju na 3' kraju mRNA. Introni se izrezuju iz primarnog RNA transkripta, a egzoni se spajaju u zrelu, funkcionalnu mRNA koja se translatira u protein. Geni viših eukariota imaju puno veću količinu DNA u intronskim sekvencama nego u egzonima. Najveći broj introna pripada klasi *spliceosomalnih* introna čije je izrezivanje katalizirano velikim proteinskim kompleksom *spliceosomom*. (Barbazuk *et al.*, 2008) *Spliceosom* se sastoji od specijaliziranih RNA-proteinskih kompleksa, snRNPs (*small nuclear ribonucleoproteins*) od kojih svaki sadrži snRNA (U1, U2, U3, U4, U5, U6). (Anireddy *et al.*, 2012) *Spliceosomalni* introni viših eukariota najčešće (u 99% slučajeva) imaju dinukleotidnu sekvencu GU na 5' kraju (donorsko mjesto) i AG na 3' kraju (akceptorsko mjesto). (Anireddy *et al.*, 2012) snRNA U1 sadrži sekvencu komplementarnu sekvencama na 5' kraju introna i veže se na tu regiju u primarnom transkriptu. Vezanje snRNPs U2, U4, U5 i U6 dovodi do formiranja *spliceosoma*. Aktivni kompleks (snRNPs U2, U5 i U6) katalizira dvije transesterifikacije prilikom čega se izreže intron i dođe do ligacije dvaju egzona. Intron ostaje u jezgri i razgrađuje se. (Nelson, D. L., Cox, M. M., 2008)

Eukarioti proizvode različite oblike zrele mRNA (izoforme) iz jednog genskog transkripta procesom alternativnog izrezivanja introna (*alternative splicing*). Ukoliko dva ili više zrelih mRNA transkriptata dobivenih iz istoga lokusa imaju različite strukture, pretpostavlja se da je došlo do alternativnog izrezivanja introna. Da bi se moglo detektirati alternativno izrezivanje introna, mora biti poznata sekvenca zrele mRNA i originalna genomska sekvenca. Alternativno izrezivanje introna detektira se pomoću ESTs (*expressed sequence tags*) usklađenih s genomom tako da se traže strukturalne različitosti, pomoću RT-PCR/RACE (*reverse transcriptase-mediated PCR*) za specifični lokus i traženjem izoformi te pomoću *microarray* metoda koje pokrivaju lokus i služe za detekciju izoformi. (Anireddy *et al.*, 2012) Opsežno EST sekvencioniranje pokazalo je da je alternativno izrezivanje introna široko rasprostranjeno među višim eukariotima i da uvelike utječe na kompleksnost njihova transkriptoma; više od 60%

ljudskih gena i 20 – 30% biljnih gena prolaze alternativno izrezivanje introna. (Severing *et al.*, 2009)

Na donorskom mjestu introna (32 pb oko 5' kraja introna) nalazi se konsenzusna sekvenca od polipirimidinskog nizvodnog slijeda i GT dinukleotida na mjestu izrezivanja introna, a na akceptorskom se mjestu (32 pb od 3' kraja introna) također nalazi konsenzusna sekvenca od polipirimidinskog uzvodnog slijeda i dinukleotida AG na mjestu izrezivanja introna. (Anireddy *et al.*, 2012) Analizama alternativnog izrezivanja introna ustanovljeno je nekoliko mogućnosti za nastanak izoformi mRNA (slika 1): alternativno donorsko mjesto, alternativno akceptorsko mjesto, alternativni terminalni egzon, preskočeni egzon, alternativna inicijacija unutar introna, alternativna terminacija unutar introna, neizrezani intron. (Barbazuk *et al.*, 2008)



Slika 1. Mogućnosti nastanka izoformi mRNA alternativnim izrezivanjem introna (prema Barbazuk *et al.*, 2008.)

Alternativno izrezivanje introna široko je rasprostranjen fenomen u višim eukariotima, no još je upitno u kojoj mjeri dovodi do funkcionalnih izoformi proteina i

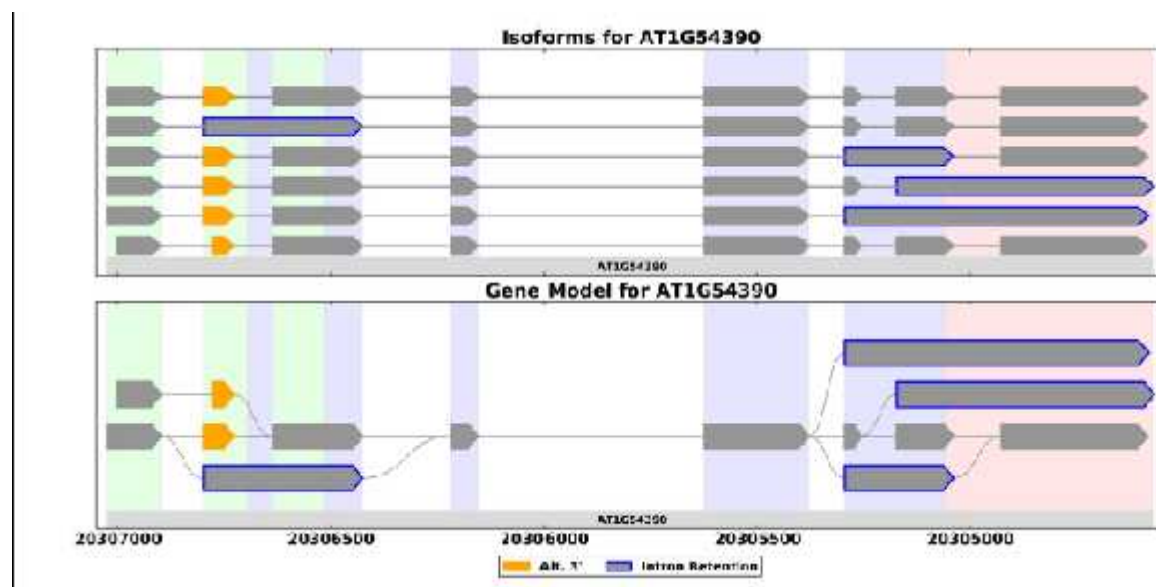
proširenja proteoma. Dobro je istražen kod životinjskih vrsta, no kod biljaka se još uvijek ne zna njegov puni značaj i očekuje se puno istraživanja, sekvencioniranja RNA biljaka od interesa radi dobivanja šire slike o alternativnom izrezivanju introna u biljaka. Glavno pitanje koje se postavlja jest služi li alternativno izrezivanje introna u biljkama samo kao doprinos proteomskoj raznolikosti putem proizvodnje funkcionalnih izoformi proteina ili ima i važnu ulogu u posttranskripcijskoj regulaciji ekspresije gena putem proizvodnje izoformi zrelih mRNA koje sadrže prerani terminacijski kodon, što aktivira degradaciju transkripta pomoću NMD (*nonsense mediated decay*) putova razgradnje mRNA. Kao modeli za analizu alternativnog izrezivanja introna najviše se koriste od dvosupnica uročnjak (*Arabidopsis thaliana*) i od jednosupnica riža (*Oryza sativa*). Koristeći se podacima iz odabranih relevantnih članaka, pružila sam kratak pregled poznatih informacija o procesima alternativnog izrezivanja introna u biljkama. U sklopu praktičnoga dijela ovog seminarskog rada proučila sam sekvence genomske DNA skupine gena *BPM* iz *A. thaliana* i predvidjela neka moguća alternativna mjesta izrezivanja introna u tim genima te aminokiselinski sastav proteina.

2. ALTERNATIVNO IZREZIVANJE INTRONA U BILJAKA

Oko 20 – 30% biljnih gena prolazi alternativno izrezivanje introna. (Severing *et al.*, 2009) Kod cvjetnica 48% gena koji sadrže introne prolazi kroz te procese. (Anireddy *et al.*, 2012) Komparativne analize biljaka i životinja pokazale su velike razlike u procesima alternativnog izrezivanja introna te su pokazale da biljni i životinjski organizmi na različite načine prepoznaju egzone i introne. Također su uočene i razlike u učestalosti procesa alternativnog izrezivanja introna između biljaka i životinja. Za biljke postoji puno manje provedenih istraživanja te se razvijaju nove metode za identifikaciju cis-elemenata uključenih u regulaciji alternativnog izrezivanja introna. Biljni geni uglavnom su kraći nego životinjski, sadrže manje egzona i puno kraće introne (npr. kod uročnjaka je prosječna dužina egzona 173, a introna 172 nukleotida; kod riže su egzoni prosječno dugi 193, a introni 433 nukleotida, dok, npr., introni ljudskih gena u prosjeku imaju 3000 nukleotida). (Anireddy *et al.*, 2012) Većina istraživanja za cilj ima identifikaciju elemenata biljnih genoma uključenih u regulaciju procesa alternativnog izrezivanja introna te se zapravo pokušava dešifrirati biljni *splicing* kod.

Odnedavno dostupni podaci o sekvencama biljnoga genoma i samim time dostupni podaci o transkriptima omogućili su globalnu analizu procesa alternativnog izrezivanja introna kod mnogih biljnih vrsta. Godine 2001. bilo je poznato samo 36 biljnih gena koji prolaze procese alternativnog izrezivanja introna. (Anireddy *et al.*, 2012) No, sekvenciranje cijelog genoma uročnjaka (u novije vrijeme i genoma drugih biljnih vrsta), kao i dostupnost ogromnih količina podataka o transkribiranim sekvencama u obliku EST-a (*expressed sequence tags*) i cDNA, te ograničena količina podataka dobivenih NGS (*next generation sequencing*) tehnikama omogućilo je analizu procesa alternativnog izrezivanja introna u cijelom transkriptomu nekoliko biljnih vrsta, uključujući uročnjak, rižu, grožđe i krastavac. Cjelovite cDNA (*full-length cDNAs*) pružaju najbolju potvrdu strukture za različite izoforme transkripata dobivene izrezivanjem introna. Usklađujući takvu sekvencu (cDNA) s referentnim genomom, dobiva se uvid u točnu egzon-intron strukturu transkripta. EST-ovi (nasumični uzorci mRNA stanice, nejednoliko pokrivaju gene) su kraći, ali svejedno pokrivaju nekoliko egzona. NGS očitavanja su kratka (oko 100 pb) i ona pružaju dokaze samo lokalnih struktura transkripta. Većina metoda za sklapanje transkriptoma podrazumijeva konstrukciju kompaktnog grafičkog prikaza egzon-intron strukture gena koji uključuje sve načine na

koje egzoni određenoga gena mogu biti sklopljeni u transkript; taj prikaz je tzv. *splice graph* (slika 2). Kod biljaka do 56% izoformi transkripata nastalih alternativnim izrezivanjem introna nastane zbog neizrezanog introna. Kod životinja izoforme transkripata najčešće nastaju preskakanjem egzona (58% kod ljudi), dok je npr. kod uročnjaka to uzrok nastanku manje od 8% izoformi transkripata. (Anireddy *et al.*, 2012)



Slika 2. Izoforme transkripata dobivenih izrezivanjem introna za gen *ING2* iz uročnjaka *A. thaliana* prikazane kao set transkripata i odgovarajući *splice graph* (prema Anireddy *et al.*, 2012.)

Budući da su biljni introni kraći od životinjskih te da je kod biljaka najčešći oblik alternativnog izrezivanja introna upravo neizrezivanje introna (tzv. zaostali intron), pretpostavlja se da je za izrezivanje introna u biljaka ključan neidentificirani mehanizam prepoznavanja introna. Biljni introni bogati su U i UA nukleotidima, a biljni egzoni bogati su G nukleotidima. (Wang i Brendel, 2005) Nekoliko istraživanja upućuje na to da su cis-elementi uključeni u prepoznavanje introna kod biljaka različiti od cis-elemenata kod kvasaca i životinja. (Barbazuk *et al.*, 2008) Ta pretpostavka utemeljena je na činjenici da životinjski introni ne budu točno izrezani u biljkama i obrnuto. Visok sadržaj U i UA nukleotida u biljnim intronima pokazao se važnim za prepoznavanje mjesta izrezivanja introna i njihovo efikasno izrezivanje kod introna U2 tipa, što upućuje na prisutstvo proteina koji ostvaruju interakcije s U ili UA bogatim elementima. Istraživanja su pokazala da do alternativnog izrezivanja introna najčešće dolazi kod introna sa smanjenim sadržajem UA nukleotida. No, premalo je provedenih istraživanja na ovom području kako bi postojale preciznije spoznaje o nekim cis-elementima koji bi potvrdili

da je neki određeni mehanizam prepoznavanja introna ključan za pravilno izrezivanje introna, tj. zaostajanje introna. (Anireddy *et al.*, 2012)

Uz proširenje proteoma funkcionalnim izoformama proteina alternativno izrezivanje introna može poslužiti kao posttranskripcijski mehanizam za regulaciju ekspresije gena putem proizvodnje izoformi mRNA koje sadrže prerani terminacijski kodon koji aktivira degradaciju transkripta NMD putem razgradnje mRNA. Iako većina alternativnog izrezivanja introna ima potencijal za proizvodnju različitih proteinskih izoformi, nije poznato u kojoj su mjeri proizvedene izoforme funkcionalne. Brojna istraživanja u kojima su prikazane strukture proteina dobivenih alternativnim izrezivanjem introna pružila su uvid u utjecaj alternativnog izrezivanja introna na stabilnost i funkciju proteina. No, informacije o strukturi dostupne su samo za mali dio poznatih proteina. Alternativni pristup je usporedna analiza alternativnog izrezivanja introna dviju ili više vrsta. Osnovna je pretpostavka za taj pristup tvrdnja da je u vrstama vjerojatnija konzerviranost funkcionalnih genetskih značajki nego onih nefunkcionalnih. Te značajke mogu biti genski specifične, poput egzon-intron strukture određenoga gena, ili povezane s određenim staničnim procesima. Kad je riječ o konzerviranosti alternativnog izrezivanja introna, to znači potragu unutar ortolognih gena (homologni geni koji dijele istu funkciju, makar se danas mogu naći u vrlo različitim organizmima) za sličnim procesima alternativnog izrezivanja introna koji su vjerojatno bili prisutni i u zajedničkom pretku proučavanih vrsta. (Severing *et al.*, 2009) Pretpostavka je da su konzervirani jer su predstavljali dugotrajnu selekcijsku prednost. Na razini strukture gena, konzervirani procesi alternativnog izrezivanja introna definirani su kao slični procesi na ortolognim intronima/egzonima barem dviju vrsta. No, konzervirani procesi alternativnog izrezivanja introna nemaju nužno iste učinke na proteine koji su kodirani određenim genima, dok nekonzervirani procesi alternativnog izrezivanja introna mogu proizvesti iste učinke na protein. Istraživanjima na različitim vrstama koja su provedena tijekom nekoliko posljednjih godina otkriveno je nekoliko općih značajki alternativnog izrezivanja introna. Na primjer, frakcije svih gena koji prolaze alternativno izrezivanje introna kod *Arabidopsis thaliana* i *Oryza sativa* su slične, a slična je i učestalost individualnih procesa alternativnog izrezivanja introna u objema vrstama. Komparativna analiza gena širokog spektra koji su prošli alternativno izrezivanje introna pokazala je da geni koji kodiraju za proteine određenih funkcija (npr. DNA-, RNA-, Ca-vezni proteini) imaju povišene razine alternativnog izrezivanja introna.

Uspoređivano je alternativno izrezivanje introna u različitim kategorijama gena prema *GeneOntology* podjeli. Geni za ribosomalne proteine i geni uključeni u prijenos signala većinom imaju manje izoformi mRNA, dok geni uključeni u replikaciju DNA i stanični ciklus uglavnom imaju više izoformi. Također je pokazano da geni koji kodiraju za proteine koji ulaze u interakciju s drugim proteinima imaju veću učestalost alternativnog izrezivanja introna od gena koji kodiraju za proteine koji ne ostvaruju međuproteinske interakcije. (Severing *et al.*, 2009) Pri komparativnoj analizi, alternativno izrezivanje introna uočeno je u 24% testiranih lokusa uročnjaka i u 28% testiranih lokusa riže. Kod kukuruza je uočeno alternativno izrezivanje introna u 40% testiranih lokusa, ali distribucija procesa alternativnog izrezivanja introna slična je kao i kod uročnjaka i riže. Nakon konstrukcije hipotetskih izoformi proteina dobivenih alternativnim izrezivanjem introna za uročnjak dobiveno je 36 950 izoformi, za rižu 40 543 i za kukuruz 25 064. Konstruirana su dva seta procesa alternativnog izrezivanja introna ortolognih parova:

a) ortologni parovi s procesima alternativnog izrezivanja introna koji su pretpostavljeni kao mete za NMD put razgradnje u objema vrstama

b) ortologni parovi s procesima alternativnog izrezivanja introna za koje se pretpostavlja da će se translatirati u proteine u objema vrstama (njih je bilo dvostruko više).

Primijećen je samo mali broj slučajeva u kojima su procesi alternativnog izrezivanja introna u oba ortologa rezultirali u sličnim modifikacijama u odnosu na originalni proteinski produkt. (Severing *et al.*, 2009) Alternativnim izrezivanjem introna kojim nastaje funkcionalna izoforma proteina gen postaje polimorfan u svojoj funkciji. Kada i originalni proteinski produkt i izoforma nastala alternativnim izrezivanjem introna imaju određenu selektivnu prednost, oba se mogu zadržati tijekom evolucije. Čak i procesi alternativnog izrezivanja introna koji vode u NMD put razgradnje mRNA mogu predstavljati određenu selektivnu prednost. Rezultati istraživanja pokazali su da su frakcije ortolognih parova koje su prošle iste procese alternativnog izrezivanja introna na različitim mjestima i onih koji su prošli različite procese alternativnog izrezivanja introna na istim mjestima poprilično slične i u a) setu ortolognih i u b) setu ortolognih parova. S obzirom da izoforme iz a) seta ortolognih parova nikada neće funkcionirati kao proteini, ta sličnost upućuje na to da određena količina

konzerviranosti procesa alternativnog izrezivanja introna nije rezultat funkcionalne konzerviranosti kroz evoluciju. Dakle, utvrđeno je da konzerviranost procesa alternativnog izrezivanja introna nije rezultat očuvanja funkcionalnog proteinskog polimorfizma potaknutog alternativnim izrezivanjem introna. (Severing *et al.*, 2009) Računalnim analizama genoma i proučavanjem funkcionalnih proteinskih domena kodiranih genima koji su prošli normalno i alternativno izrezivanje introna otkriveno je da je samo RRM (*RNA recognition motif*) pretjerano zastupljen u proteinima dobivenim alternativnim izrezivanjem introna u svim analiziranim vrstama. Tri tipa domena najviše su zastupljena u proteinima translatiranim od transkripata dobivenih normalnim izrezivanjem introna. Analiza procesa alternativnog izrezivanja introna koji dovode do uklanjanja cijelih proteinskih domena otkrila je da je samo mali broj tipova domena izrezan u svim analiziranim vrstama. U slučajevima kada i jest cijela proteinska domena izrezana, ona je bila dio tandemskog ponavljanja. (Severing *et al.*, 2009)

Kako raste dostupnost biljnih genomskih sekvenci, raste i broj komparativnih analiza koje će identificirati važne procese alternativnog izrezivanja introna u biljkama, otkriti utjecaj duplikacije genoma na evoluciju alternativnog izrezivanja introna i otkriti specifične cis-elemente koji reguliraju te procese u biljkama. Pretpostavlja se da su procesi alternativnog izrezivanja introna u biljkama uključeni u važne životne funkcije (poput odgovora na stresne podražaje) te da mogu imati utjecaj na domestikaciju bilja i selekciju svojstava.

3. CILJ RADA

Geni porodice *BPM* kodiraju za proteine s MATH i BTB domenama koji se vežu na druge proteine (npr. transkripcijske regulatore) te na taj način reguliraju funkciju drugih proteina. Također, BPM proteini sudjeluju u prepoznavanju proteina koje treba ubikvitinirati. Takvi ubikvitinirani proteini mogu se degradirati putem proteasoma, ili ubikvitinacija može utjecati na njihovu staničnu lokalizaciju i interakcije s drugim proteinima i/ili nukleinskim kiselinama te tako utjecati na njihovu funkciju.

Alternativnim izrezivanjem introna *BPM* gena teoretski je moguće dobiti subvarijante BPM proteina koje imaju drugačiju specifičnost za ciljne proteine.

Uročnjak ima 6 *BPM* gena i do sada je poznato 9 varijanti mRNA koje se prevode u proteine. Ista porodica gena u riži sastoji se od 71 gena koji većinom ne sadrže introne. Neki od rižinih gena ove porodice aktiviraju se razvojno specifično.

U ovom će radu biti predviđene varijante alternativnog izrezivanja introna gena *BPM* iz uročnjaka. Dobivene alternativne varijante bit će uspoređene s rižinim genima.

4. MATERIJALI I METODE

U ovom će radu biti predviđene varijante gena *BPM* uročnjaka.

Sekvence genomskih DNA *BPM* gena uročnjaka preuzela sam iz baze podataka TAIR (*The Arabidopsis Information Resource*). Geni *BPM* nalaze se na sljedećim lokusima:

BPM1: AT5G19000

BPM2: AT3G06190

BPM3: AT2G39760

BPM4: AT3G03740

BPM5: AT5G21010

BPM6: AT3G43700

Za prevođenje nukleotidnih sekvenci u proteinske koristila sam program ExPASy (*Expert Protein Analysis System*) portala Švicarskog instituta za bioinformatiku (SIB).

Varijante alternativnog izrezivanja introna predvidjela sam na temelju građe *spliceosomalnih* introna viših eukariota najčešće. U pravilu (u 99% slučajeva) introni imaju dinukleotidnu sekvencu GU na 5' kraju (donorsko mjesto) te AG na 3' kraju (akceptorsko mjesto). ([Anireddy et al., 2012](#))

Dobivene varijante proteina BPM usporedit ću s poznatim EST sekvencama pomoću programa BLAST (The **B**asic **L**ocal **A**lignment **S**earch **T**ool, <http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

5. REZULTATI

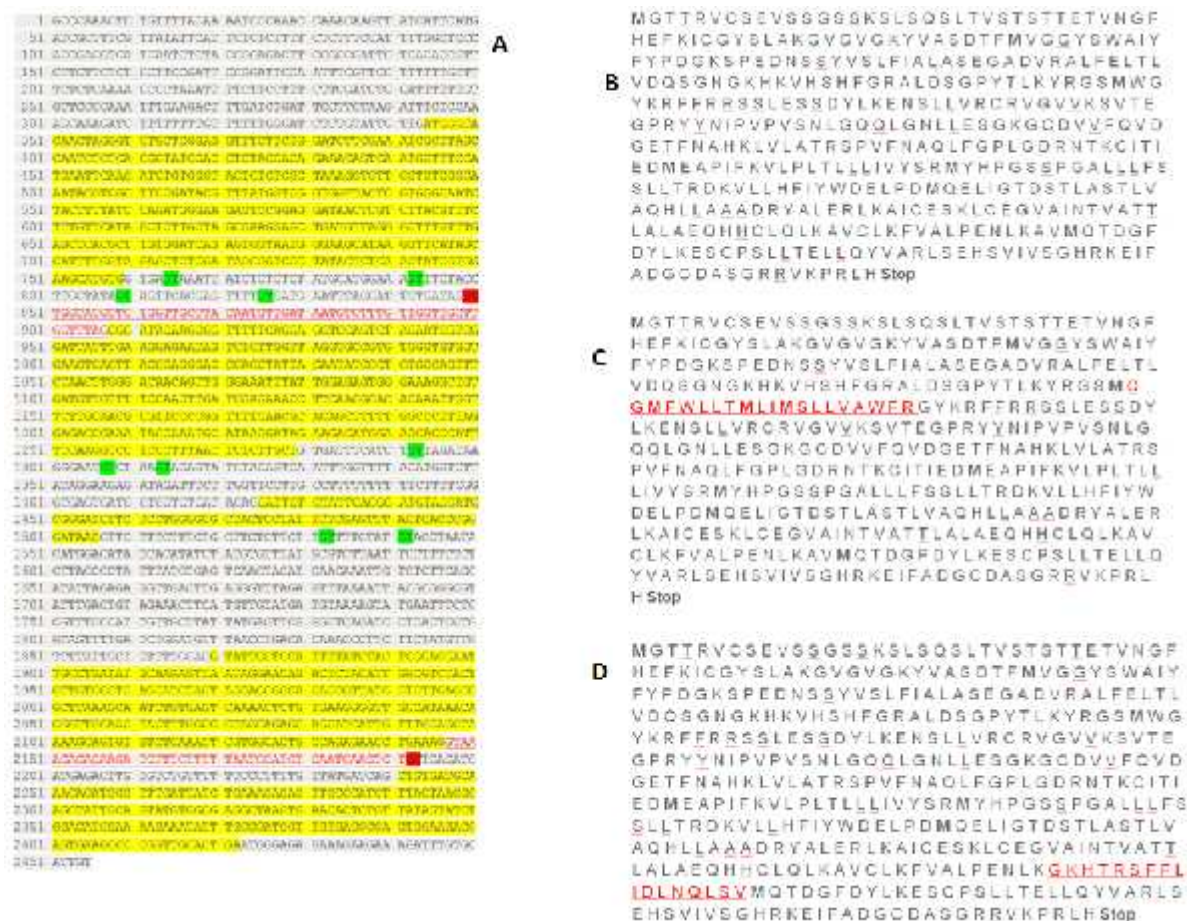
Alternativni oblici proteina prikazani u ovom dijelu rada (tj. proteini koji bi nastali kao posljedica alternativnog izrezivanja introna na mjestima koja sam predviđela gledajući sekvence genomskih DNA *BPM* gena uročnjaka) rezultat su samo mojih predviđanja, postojanje tih oblika nije ni na koji način eksperimentalno potvrđeno. Tijekom predviđanja varijanti alternativnog izrezivanja introna (odabirala sam moguća alternativna akceptorska i donorska mjesta, unosila hipotetske egzonske nukleotidne sekvence u program ExPASy) kod prevođenja nukleotidnih sekvenci u proteinske dobivala sam mnoge varijante s nizom terminacijskih kodona zbog kojih takve varijante ne bi bile funkcionalni proteini. Takva „neodgovarajuća“ alternativna akceptorska i donorska mjesta označila sam na svim slikama genomskih DNA proučavanih gena, a takve sekvence mogle bi predstavljati supstrate za NMD put razgradnje.

Dosad su u banci gena prikazane dvije alternativne varijante za transkript *BPM1* (slike 3. i 4.). Gen *BPM1* mogao bi, prema nukleotidnom sastavu, uz opisane imati i alternativne varijante mRNA. Ukoliko bi do alternativnog izrezivanja introna došlo na akceptorskom mjestu na 802. pb, nastaje protein koji ima 20 aminokiselina više (slika 3. C). Ukoliko bi došlo do alternativnog izrezivanja introna korištenjem alternativnog donorskog mjesta na 2145. pb, nastaje protein koji ima 17 aminokiselina više (slika 3. D).

[illegible]

Slika 3. Sekvenca genomske DNA gena BPM1, žuto označeni dijelovi – egzoni, **AG** – alternativno akceptorsko mjesto, **GT** – alternativno donorsko mjesto, **GT** – hipotetska alternativna donorska mjesta, no prevedene proteinske sekvence bi imale niz STOP kodona (A); normalni protein, dobiven uobičajenim *splicingom* (B); protein dobiven od alternativne izoforme mRNA zbog alternativnog akceptorskog mjesta (C); protein dobiven od alternativne izoforme mRNA zbog alternativnog donorskog mjesta (D)

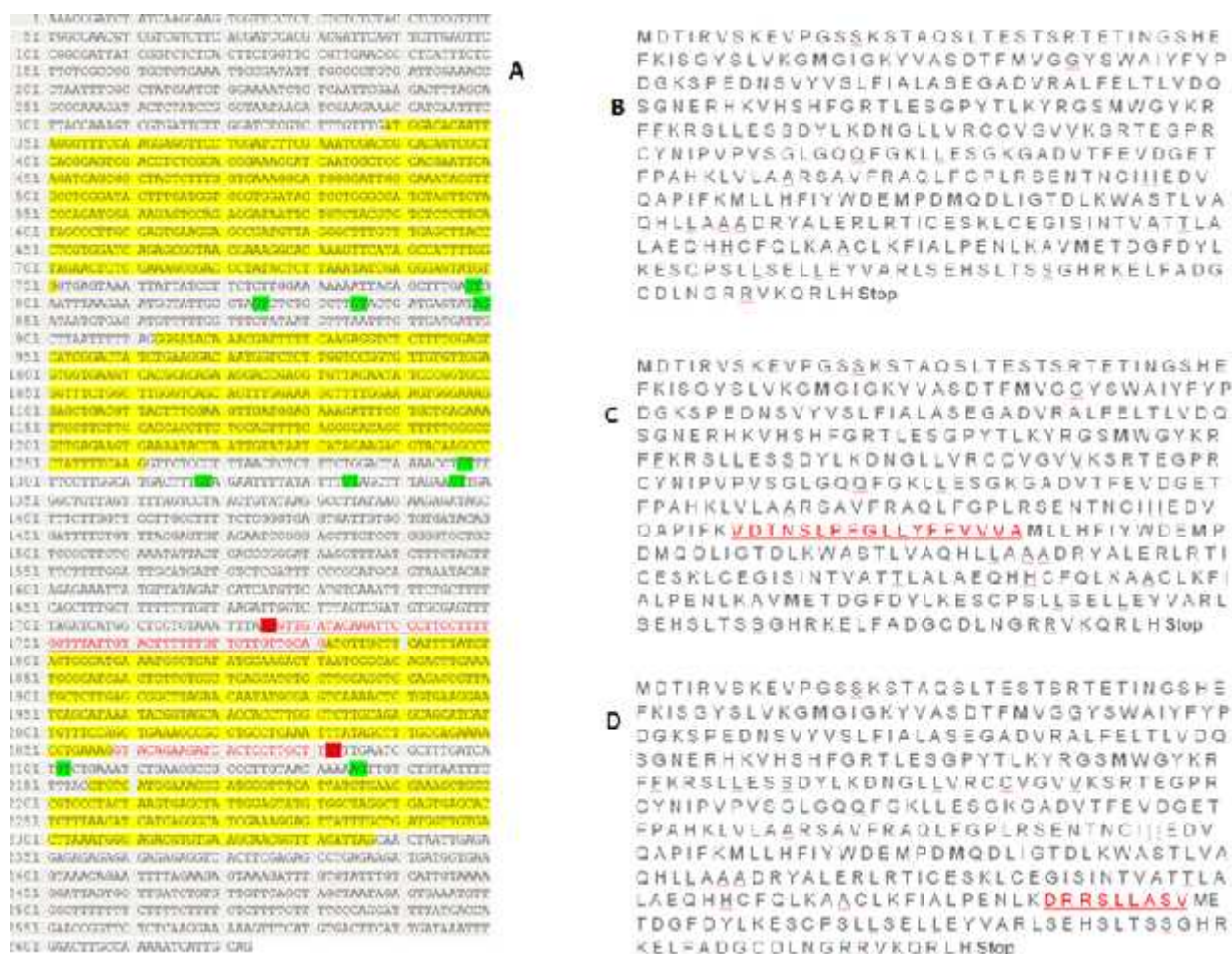
Izoforma 2 *BPM1* gena ima dodatni egzon te sadrži 35 aminokiselina više od izoforme 1. Ako bi kod druge varijante *BPM1* do alternativnog izrezivanja introna došlo zbog alternativnog akceptorskog mjesta na 849. pb, nastao bi protein koji ima 21 aminokiselinu više (slika 4. C), a korištenjem alternativnog donorskog mjesta na 2192. pb nastaje protein sa 17 aminokiselina više (slika 4. D).



Slika 4. Sekvenca genomske DNA gena BPM1, žuto označeni dijelovi – egzoni, **AC** – alternativno akceptorsko mjesto, **GT** – alternativno donorsko mjesto, **CT** – hipotetska alternativna donorska mjesta, no prevedene proteinske sekvence bi imale niz STOP kodona (A); normalni protein dobiven normalnim *splicingom* (B); protein dobiven od alternativne izoforme mRNA zbog alternativnog akceptorskog mjesta (C); protein dobiven od alternativne izoforme mRNA zbog alternativnog donorskog mjesta (D)

Dosad su u banci gena prikazane dvije varijante za transkript *BPM2* (slika 5. i 6.). Izoforma 2 *BPM2* gena ima različite i kraće egzone 3 i 4 te je kraća za 111 aminokiselina od izoforme 1. Ako bi kod prve varijante *BPM2* do alternativnog izrezivanja introna

došlo zbog alternativnog akceptorskog mjesta na 1725. pb, nastaje protein koji ima 18 aminokiselina više (slika 5. C), a korištenjem alternativnog donorskog mjesta na 2081. pb nastaje protein s 9 aminokiselina više (slika 5. D).



Slika 5. Sekvenca genomske DNA gena BPM2, žuto označeni dijelovi – egzoni, **AC** – alternativno akceptorsko mjesto, **GT** – alternativno donorsko mjesto, **AG**, **GT** – hipotetska alternativna donorska i akceptorska mjesta, no prevedene proteinske sekvence imale bi niz STOP kodona (A); normalni protein dobiven normalnim *splicingom* (B); protein dobiven od alternativne izoforme mRNA zbog alternativnog akceptorskog mjesta (C); protein dobiven od alternativne izoforme mRNA zbog alternativnog donorskog mjesta (D)

Ako bi do alternativnog izrezivanja introna kod druge varijante *BPM2* došlo zbog alternativnog akceptorskog mjesta na 2135. pb, nastao bi protein koji ima 18 aminokiselina više (slika 6. C). Ukoliko u obzir uzmem alternativno akceptorsko mjesto na 2154. pb, dobivam protein za koji sam utvrdila da zaista postoji uspoređujući tu

A

MDTIRVSKVEVPGSSSKSTAQSLTESTSRTETINGSHE
FKISGYSLVKGMGIGKYVASDTFMVGGYSWAIYFYP
DGKSPEDNSVYVSLFIALASEGADVRLFELTVDDQ
SGNERHKVHSHFGRTLESSGPYTLKYRGSWMWGYKR
FEKRSLLLESSDYLDKNGLLVRCQVGVYKSRTEGPR
CYNIPVPVSGLGQFGKLLSGLKGAADVTFEVDGET
FPAHKLVLAAARS AVFRAQLFGPLRSNTNSLEVEAE
SCPSLLSELLEYVARLSEHSLTSSGHRKELFADGCD
LNGRRVYKQRLH Stop

B

MDTIRVSKVEVPGSSSKSTAQSLTESTSRTETINGSHE
FKISGYSLVKGMGIGKYVASDTFMVGGYSWAIYFYP
DGKSPEDNSVYVSLFIALASEGADVRLFELTVDDQ
SGNERHKVHSHFGRTLESSGPYTLKYRGSWMWGYKR
FEKRSLLLESSDYLDKNGLLVRCQVGVYKSRTEGPR
CYNIPVPVSGLGQFGKLLSGLKGAADVTFEVDGET
FPAHKLVLAAARS AVFRAQLFGPLRSNTNSLEVEAE
SCPSLLSELLEYVARLSEHSLTSSGHRKELFADGCD
LNGRRVYKQRLH Stop

C

MDTIRVSKVEVPGSSSKSTAQSLTESTSRTETINGSHE
FKISGYSLVKGMGIGKYVASDTFMVGGYSWAIYFYP
DGKSPEDNSVYVSLFIALASEGADVRLFELTVDDQ
SGNERHKVHSHFGRTLESSGPYTLKYRGSWMWGYKR
FEKRSLLLESSDYLDKNGLLVRCQVGVYKSRTEGPR
CYNIPVPVSGLGQFGKLLSGLKGAADVTFEVDGET
FPAHKLVLAAARS AVFRAQLFGPLRSNTNSLEVEAE
SVISLAYMETDGFYKLESCPSLLSELLEYVARLSE
HSLTSSGHRKELFADGCDLNGRRVYKQRLH Stop

D

MDTIRVSKVEVPGSSSKSTAQSLTESTSRTETINGSHE
FKISGYSLVKGMGIGKYVASDTFMVGGYSWAIYFYP
DGKSPEDNSVYVSLFIALASEGADVRLFELTVDDQ
SGNERHKVHSHFGRTLESSGPYTLKYRGSWMWGYKR
FEKRSLLLESSDYLDKNGLLVRCQVGVYKSRTEGPR
CYNIPVPVSGLGQFGKLLSGLKGAADVTFEVDGET
FPAHKLVLAAARS AVFRAQLFGPLRSNTNSLEVEAE
VMDTGFYKLESCPSLLSELLEYVARLSEHSLTSS
GHRKELFADGCDLNGRRVYKQRLH Stop

18

A

Phylogenetic tree showing the relationships between 100 sequences. The tree is rooted at the bottom left and branches out to the right. The sequences are color-coded: green (top clade), red (middle clade), blue (bottom clade), and yellow (outgroup). A scale bar of 0.1 is shown at the bottom left.

B

Amino acid sequence of the VEGF family. The sequence is shown in a single line, with color-coded residues corresponding to the clades in Panel A. The sequence is: M E T V G G I E Q L I P D S V S T S F I E T V N G S H Q F T I Q G Y S L A K G M S P G K F I Q S D I F S V G Q Y D W A I Y F Y P D G K N P E D Q S S Y I S L F I A L A S D S N D I R A L F E L T L M D Q S G K G K H K V H S F D R A L E G G P Y T L K Y K G S M W G Y K R F E K R S A L E T S O Y L K D D C L V I N C T V G V R A R L E G P K C Y G I V L P L S N M G G I K D L D S E V G C D I A F Q V G D E T Y K A H K L I A A R S P V F R A Q F E G P I G N N N V D R I V I D I E P S I F K A M L S F I Y T D V L P N V H E I T G S S A S S F T N M I Q H L L A A A D L Y D L A R L K I L C E V L L C E K L D V D N V A T I L A L A E Q H Q F L K A F C L E F V A S P A N L G A V M K S E G F K H L K Q S G P T L L S E L L N T V A A A D K S S T S G Q S N K K R S A S S V L G C D T T N V R Q L R R R R T R K E V R A V S Stop

C

Amino acid sequence of the VEGF family. The sequence is shown in a single line, with color-coded residues corresponding to the clades in Panel A. The sequence is: M E T V G G I E Q L I P D S V S T S F I E T V N G S H Q F T I Q G Y S L A K G M S P G K F I Q S D I F S V G Q Y D W A I Y F Y P D G K N P E D Q S S Y I S L F I A L A S D S N D I R A L F E L T L M D Q S G K G K H K V H S F D R A L E G G P Y T L K Y K G S M W G Y K R F E K R S A L E T S O Y L K D D C L V I N C T V G V R A R L E G P K C Y G I V L P L S N M G G I K D L D S E V G C D I A F Q V G D E T Y K A H K L I A A R S P V F R A Q F E G P I G N N N V D R I V I D I E P S I F K A M L S F I Y T D V L P N V H E I T G S S A S S F T N M I Q H L L A A A D L Y D L A R L K I L C E V L L C E K L D V D N V A T I L A L A E Q H Q F L K A F C L E F V A S P A N L G A V M K S E G F K H L K Q S G P T L L S E L L N T V A A A D K S S T S G Q S N K K R S A S S V L G C D T T N V R Q L R R R R T R K E V R A V S Stop

D

Amino acid sequence of the VEGF family. The sequence is shown in a single line, with color-coded residues corresponding to the clades in Panel A. The sequence is: M E T V G G I E Q L I P D S V S T S F I E T V N G S H Q F T I Q G Y S L A K G M S P G K F I Q S D I F S V G Q Y D W A I Y F Y P D G K N P E D Q S S Y I S L F I A L A S D S N D I R A L F E L T L M D Q S G K G K H K V H S F D R A L E G G P Y T L K Y K G S M W G Y K R F E K R S A L E T S O Y L K D D C L V I N C T V G V R A R L E G P K C Y G I V L P L S N M G G I K D L D S E V G C D I A F Q V G D E T Y K A H K L I A A R S P V F R A Q F E G P I G N N N V D R I V I D I E P S I F K A M L S F I Y T D V L P N V H E I T G S S A S S F T N M I Q H L L A A A D L Y D L A R L K I L C E V L L C E K L D V D N V A T I L A L A E Q H Q F L K A F C L E F V A S P A N L G A V M K S E G F K H L K Q S G P T L L S E L L N T V A A A D K S S T S G Q S N K K R S A S S V L G C D T T N V R Q L R R R R T R K E V R A V S Stop

E

Amino acid sequence of the VEGF family. The sequence is shown in a single line, with color-coded residues corresponding to the clades in Panel A. The sequence is: M E T V G G I E Q L I P D S V S T S F I E T V N G S H Q F T I Q G Y S L A K G M S P G K F I Q S D I F S V G Q Y D W A I Y F Y P D G K N P E D Q S S Y I S L F I A L A S D S N D I R A L F E L T L M D Q S G K G K H K V H S F D R A L E G G P Y T L K Y K G S M W G Y K R F E K R S A L E T S O Y L K D D C L V I N C T V G V R A R L E G P K C Y G I V L P L S N M G G I K D L D S E V G C D I A F Q V G D E T Y K A H K L I A A R S P V F R A Q F E G P I G N N N V D R I V I D I E P S I F K A M L S F I Y T D V L P N V H E I T G S S A S S F T N M I Q H L L A A A D L Y D L A R L K I L C E V L L C E K L D V D N V A T I L A L A E Q H Q F L K A F C L E F V A S P A N L G A V M K S E G F K H L K Q S G P T L L S E L L N T V A A A D K S S T S G Q S N K K R S A S S V L G C D T T N V R Q L R R R R T R K E V R A V S Stop

19

Ako bi kod druge varijante *BPM3* do alternativnog izrezivanja introna došlo zbog alternativnog donorskog mjesta na 1359. pb, nastaje protein koji ima 3 aminokiseline više (slika 8. C), a zbog alternativnog donorskog mjesta na 1374. pb nastaje protein sa 8 aminokiselina više (slika 8. D).

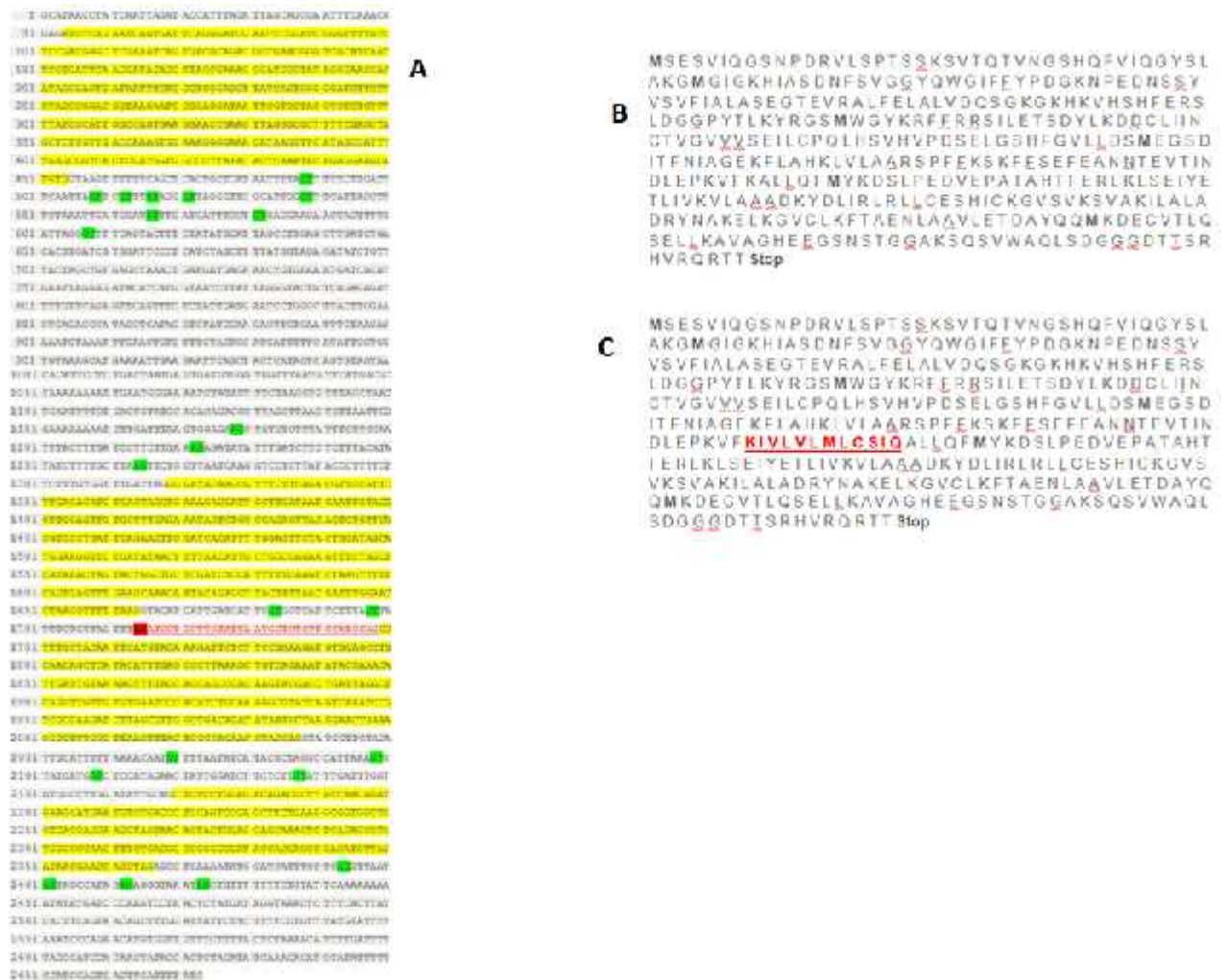


Slika 8. Sekvenca genomske DNA gena *BPM3*, žuto označeni dijelovi – egzoni, **GT** – alternativna donorska mjesta, **AG**, **GT** – hipotetska alternativna donorska i akceptorska mjesta, no prevedene proteinske sekvence imale bi niz STOP kodona (A); normalni protein dobiven normalnim *splicingom* (B); proteini dobiveni od alternativne izoforme mRNA zbog alternativnog donorskog mjesta (C, D)

[illegible]

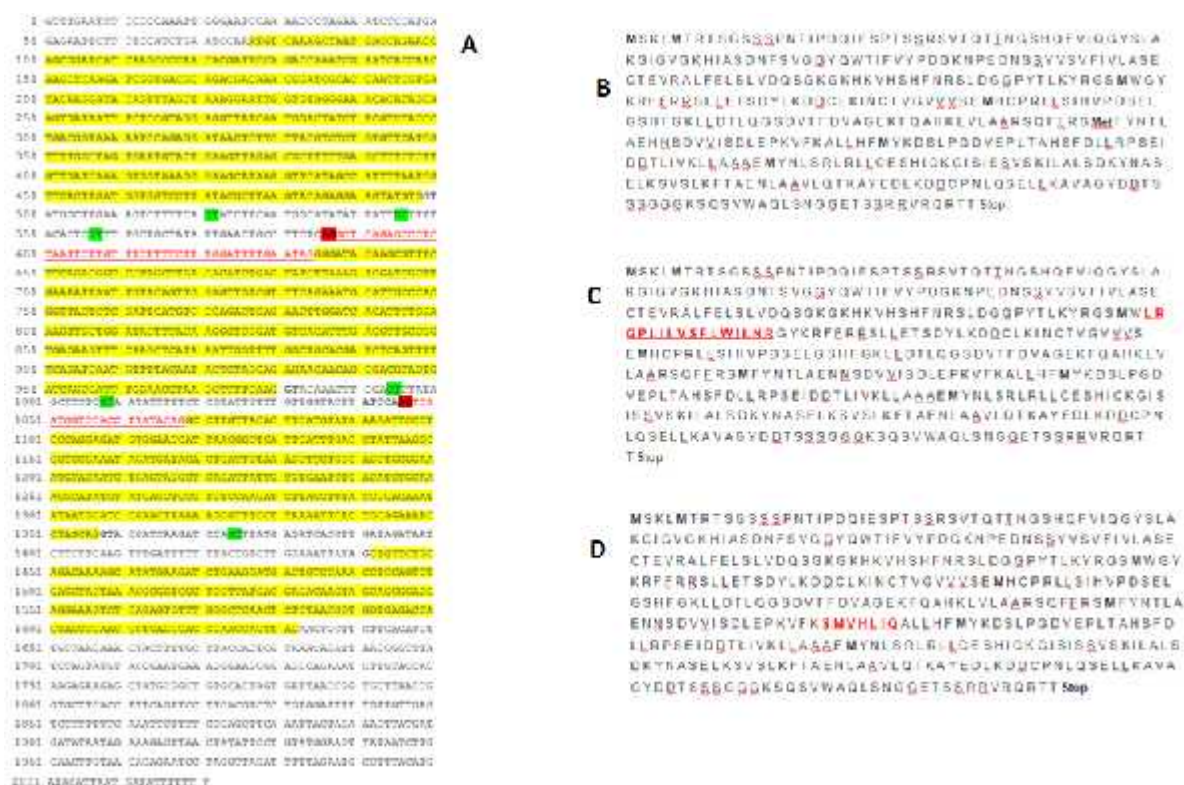
21

Ako bi kod *BPM5* došlo do alternativnog izrezivanja introna zbog akceptorskog mjesta na 1714. pb, nastaje protein sa 12 aminokiselina više (slika 10. C).



Slika 10. Sekvenca genomske DNA gena BPM5, žuto označeni dijelovi – egzoni, **AG** – alternativno akceptorsko mjesto, **AC**, **GT** – hipotetska alternativna donorska i akceptorska mjesta, no prevedene proteinske sekvence imale bi niz STOP kodona (A); normalni protein dobiven normalnim *splicingom* (B); protein dobiven od alternativne izoforme mRNA zbog alternativnog akceptorskog mjesta (C)

Ukoliko bi kod *BPM6* do alternativnog izrezivanja introna došlo zbog akceptorskog mjesta na 586. pb, nastaje protein sa 16 aminokiselina više (slika 11. C), a zbog alternativnog akceptorskog mjesta na 1046. pb nastaje protein sa 7 aminokiselina više (slika 11. D).



Slika 11. Sekvenca genomske DNA gena *BPM6*, žuto označeni dijelovi – egzoni, **AG** – alternativna akceptorska mjesta, **AG**, **GT** – hipotetska alternativna donorska i akceptorska mjesta, no prevedene proteinske sekvence imale bi niz STOP kodona (A); normalni protein dobiven normalnim *splicingom* (B); proteini dobiveni od alternativne izoforme mRNA zbog alternativnog akceptorskog mjesta (C, D)

6. ZAKLJUČAK

Alternativno izrezivanje introna dugo je smatrano relativno rijetkim događajem u eukariotskim genomima. Novija istraživanja pokazala su upravo suprotno, alternativno izrezivanje introna ima ogroman utjecaj na povećanje proteomske raznolikosti i važnu ulogu u regulaciji ekspresije gena te je pronađeno u svim višim eukariotima. Takvo odstupanje od biološke dogme „jedan gen – jedan protein“ ukazuje na potrebu za promjenama u pristupu mnogim biološkim procesima, poput proteinskih interakcija i genske ekspresije. Pretpostavlja se da više od 70% od svih gena eukariota prolazi procese alternativnog izrezivanja introna u svrhu proizvodnje funkcionalnih izoformi proteina iz jednoga gena. U zadnjem desetljeću utvrđeno je da alternativno izrezivanje introna utječe na vezna svojstva proteina, unutarstaničnu lokalizaciju, enzimatsku aktivnost, stabilnost proteina i posttranslacijske modifikacije velikog broja proteina. Posljedice alternativnog izrezivanja introna vidljive su na proteinskim alternacijama (varijante različitih proteinskih produkata) i modifikacijama na razini transkripta (varijante s različitim mogućnostima translacije i različitom stabilnosti). Iako relativno dobro istražen fenomen kod životinjskih organizama, alternativno izrezivanje introna u biljaka još uvijek nije dovoljno istraženo. Još uvijek je upitno kako točno razlikovati pogreške u izrezivanju introna od biološki funkcionalnih procesa. Prekid sekvenci koje kodiraju za proteine i/ili stvaranje preranog terminacijskog kodona ne upućuju nužno na pogreške s obzirom na to da sparivanje alternativnog izrezivanja introna i NMD putova razgradnje RNA može biti regulirano i pružati dodatnu razinu regulacije ekspresije gena. Konzerviranost procesa smatra se najboljim indikatorom biološki funkcionalnih alternativnih izrezivanja introna.

U praktičnom dijelu rada proučavala sam sekvence genomske DNA gena porodice *BPM* kod uročnjaka *A. thaliana*, te predložila neke subvarijante *BPM* proteina koji bi nastali kao rezultat alternativnog izrezivanja introna na različitim donorskim i akceptorskim mjestima. Ista porodica gena u riži sastoji se od 71 gena koji većinom ne sadrže introne, a neki se aktiviraju razvojno specifično. Predviđene varijante alternativnog izrezivanja introna gena *BPM* iz uročnjaka usporedila sam s rižinim genima pomoću programa BLAST. U slučaju izoforme 2 gena *BPM2*, ukoliko se alternativno izrezivanje introna dogodi na alternativnom akceptorskom mjestu na 2154. pb dobiva se protein za koji sam utvrdila da zaista postoji (uspoređujući tu varijantu s

postojećim proteinima koristeći program BLAST). Na slici 6. A označeno je to mjesto i EST koji postoji u banci podataka ([ref/NM_111494.2/](#)). S obzirom na to da se rižini geni bez introna aktiviraju razvojno specifično, kod uročnjaka bi procesi alternativnog izrezivanja introna mogli imati funkciju proizvodnje subvarijanata BPM proteina koje imaju drugačiju specifičnost za ciljne proteine.

Kako bi se mogle detektirati sve moguće izoforme genskih transkripata, razvijaju se novije i efikasnije metode prepoznavanja procesa alternativnog izrezivanja introna. Pretpostavke o važnosti i rasprostranjenosti alternativnog izrezivanja introna mijenjaju se iz dana u dan. Otkrivanje sve veće važnosti ovih procesa upućuje na potrebu stvaranja kataloga izoformi genskih transkripata i potrebu za potpunijim saznanjima o uzrocima, svrsi, mehanizmu i regulaciji alternativnog izrezivanja introna.

7. SAŽETAK

Alternativno izrezivanje introna proces je kojim se iz jednog genskog transkripta dobivaju različite izoforme zrele mRNA. Prisutno je u svim višim eukariotima i ima ulogu u funkcionalnom proširenju proteoma, kao i u posttranskripcijskoj regulaciji ekspresije gena. Procesi alternativnog izrezivanja introna u biljaka nisu dovoljno istraženi i tek se očekuju brojna istraživanja na tom području sukladno s razvojem novih metoda detekcije tog procesa. Sekvencioniranje genoma *A. thaliana* (a i drugih biljnih vrsta u novije vrijeme) omogućilo je nova saznanja o alternativnom izrezivanju introna u biljaka, a uročnjak je biljka koja se najčešće koristi u ovim istraživanjima. U praktičnome dijelu rada navedene su neke hipotetske izoforme proteinskih produkata *BPM* gena uročnjaka dobivene alternativnim izrezivanjem introna zbog alternativnih donorskih i akceptorskih mjesta.

SUMMARY

Alternative splicing is a process of producing different isoforms of mature mRNA which originate from a common locus. Alternative splicing is present in all higher eukaryotes and has roles in expanding proteome diversity and post-transcriptional regulation. Alternative splicing in higher plants has not been studied extensively and new studies are anticipated as the new tools to accurately predict and visualize alternative splicing are being developed. Sequencing of the *A. thaliana* genome (and of other plant species, recently) has provided new information about alternative splicing in higher plants, *Arabidopsis* being the most commonly used species in these researches. Here, in the practical part of this work, the *BPM* genes of *Arabidopsis* are studied and some of the hypothetical isoforms of protein products derived from alternative splicing using the alternative donor and acceptor splice sites are created.

ZAHVALA

Najsrdačnije zahvaljujem na stručnom vodstvu i susretljivosti *doc. dr. sc.* Nataši Bauer, svojoj mentorici.

8. LITERATURA

1. Anireddy, S. N. R. (2012). *Deciphering the plant splicing code: experimental and computational approaches for predicting alternative splicing and splicing regulatory elements*. Frontiers in plant science.
2. Barbazuk, W. B., Fu, Y., McGinnis, K. M. (2008) *Genome-wide analyses of alternative splicing in plants: Opportunities and challenges*. Genome Research. 18. str. 1381-1392.
3. Nelson, D. L., Cox, M. M. (2008) *Lehninger Principles of Biochemistry*. W. H. Freeman and Company. New York.
4. Severing, E. I. et al. (2009). *Comparative analysis indicates that alternative splicing in plants has a limited role in functional expansion of the proteome*. BMC Genomics.
5. Wang, B.-B., Brendel, V. (2005) *Genomewide comparative analysis of alternative splicing in plants*. PNAS. str. 7176-7180.
6. Zhang, X.-C., Gassmann, W. (2007) *Alternative Splicing and mRNA Levels of the Disease Resistance Gene RPS4 Are Induced during Defense Responses*. Plant Physiology. Vol. 145, str. 1577-1587.