Karakterizacija gena iz porodice p53 u dagnji Mytilus galloprovincialis

Baričević, Ana

Doctoral thesis / Disertacija

2015

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://urn.nsk.hr/urn:nbn:hr:217:451913

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-07-18**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb

Sveučilište u Zagrebu

PRIRODOSLOVNO-MATEMATIČKI FAKULTET
GEOLOŠKI ODSJEK

Ana Baričević

Karakterizacija gena iz porodice p53 u dagnji *Mytilus galloprovincialis*

DOKTORSKI RAD

Zagreb, 2015

University of Zagreb

FACULTY OF SCIENCE
DEPARTMENT OF GEOLOGY

Ana Baričević

Characterization of p53 gene family in mussel *Mytilus galloprovincialis*

DOCTORAL THESIS

Zagreb, 2015

Sveučilište u Zagrebu

PRIRODOSLOVNO-MATEMATIČKI FAKULTET
GEOLOŠKI ODSJEK

Ana Baričević

Karakterizacija gena iz porodice p53 u dagnji *Mytilus galloprovincialis*

DOKTORSKI RAD

Mentor: doc.dr.sc.Mauro Štifanić

Zagreb, 2015

University of Zagreb

FACULTY OF SCIENCE
DEPARTMENT OF GEOLOGY

Ana Baričević

Characterization of p53 gene family in mussel *Mytilus galloprovincialis*

DOCTORAL THESIS

Supervisor: doc.dr.sc. Mauro Štifanić

Zagreb, 2015

Ovaj je doktorski rad izrađen u Centru za istraživanje mora Instituta Ruđer Bošković u Rovinju, pod vodstvom doc.dr.sc. Maura Štifanića, u sklopu Sveučilišnog poslijediplomskog doktorskog studija Oceanologije pri Geološkom odsjeku Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu.

Zahvaljujem svima koji su mi pomogli pri izradi ovog rada, a posebno:

- mentoru doc.dr.sc. Mauru Štifaniću na prenesenom znanju i odličnoj suradnji u svim segmentima nastanka rada

- prof.dr.sc. Renatu Batelu jer mi je otvorio vrata znanosti i uvijek bio podrška u tom lijepom poslu

- članovima povjerenstva; prof.dr.sc. Renatu Batelu, prof.dr.sc. Goranu Klobučaru, izv.prof.dr.sc. Bojanu Hameru i dr.sc. Tvrtku Smitalu na korisnim savjetima i ispravkama koji su poboljšali ovaj rad

- Mirti, Sandi, Marinu, Emini, Ines, Jeleni, Danieli, Loreni, Mariji, Maji F. i svim ostalim kolegicama i kolegama iz Centra za istraživanje mora u Rovinju na pomoći, potpori i ugodnim druženjima

- prijateljima te svojoj maloj i velikoj obitelji jer su mi najveća sreća, a i pravi tim kad treba ☺

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Geološki odsjek

Doktorski rad

KARAKTERIZACIJA GENA IZ PORODICE p53 U DAGNJI *Mytilus galloprovincialis*

ANA BARIČEVIĆ
Centar za istraživanje mora, Institut Ruđer Bošković, G. Paliage 5, Rovinj

Geni porodice p53 (p53, p63 i p73) imaju uloge u važnim staničnim procesima kao što su zastoj staničnog ciklusa, apoptoza i procesi diferencijacije. U ovom radu identificirana su i karakterizirana četiri nepotpuna kodirajuća slijeda porodice p53 (cDNA1-4) i prvi poznati cjeloviti genomski slijed homologa porodice p53 mediteranske dagnje *Mytilus galloprovincialis*. Struktura gena homologa porodice p53 mediteranske dagnje najsličnija je genu p63 kralješnjaka te potvrđuje p63 kao ishodišni gen porodice p53 i jedini gen prisutan kod beskralješnjaka. Gen p63 dagnje *M. galloprovincialis* obuhvaća 20620 pb i organiziran je u 15 egzona povezanih sa 14 introna. Analiza je utvrdila da četiri nova kodirajuća slijeda mediteranske dagnje (cDNA1-4) kodiraju proteinske izoforme TAp63γ ili TAp63α te su i svi drugi poznati sljedovi porodice p53 mekušaca identificirani kao gen p63 ili njegove ekspresijske izoforme. Očuvana struktura gena p63 koljena Mollusca utvrđena je usporedbom strukture gena p63 mediteranske dagnje sa ostalim ortolozima, pronađenim u sekvenciranim genomima mekušaca *Aplysia californica*, *Lottia gigantea*, *Crassostrea gigas* i *Biomphalaria glabrata*. Analize su također utvrdile moguće pogreške anotacija gena p63 u sekvenciranim genomima te je time istaknuta vrijednost korištenja sravnjenja poznatih (referentnih) kodirajućih sljedova s novosekvenciranim genomskim sljedovima za pouzdanu anotaciju sekvenciranih genoma.

(108 stranica, 26 slika, 7 tablica, 122 literaturna navoda, jezik izvornika: hrvatski)

Ključne riječi: Mollusca, Mytilus, ortolog, p53, p63, p73, genom, anotacija gena

Mentor: doc.dr.sc. Mauro Štifanić

Stručno povjerenstvo za ocjenu rada:
prof.dr.sc. Renato Batel, znanstveni savjetnik
prof.dr.sc. Goran Klobučar
izv.prof.dr.sc. Bojan Hamer, viši znanstveni suradnik

Rad prihvaćen: 02.lipnja 2015. godine

University of Zagreb
Faculty of Science
Department of Geology

Doctoral thesis

CHARACTERIZATION OF p53 GENE FAMILY IN MUSSEL *Mytilus galloprovincialis*

ANA BARIČEVIĆ
Centre for Marine Research, Ruđer Bošković Institute,G. Paliaga 5, Rovinj

P53 gene family members (p53, p63 and p73) are involved in crucial cellular functions such as cell-cycle arrest, apoptosis and development. In this study four partial p53 family cDNAs (cDNA1-4) and the first complete genomic sequence of the p53 family homologue in the mollusc Mediterranean mussel *Mytilus galloprovincialis* were identified and characterized. Results indicate gene structure of *M.galloprovincialis* p53 family homologue to be most similar to vertebrate p63 gene, considering p63 as the ancestral gene of the p53 family as well as the only gene of the family present in non-chordate metazoan species. The *M.galoprovincialis* p63 gene spans through 20620 base pairs (bp) and is composed of 15 egzons interrupted with 14 introns. Analysis showed the new *M.galloprovincialis* cDNA (cDNA1-4) sequences code for TAp63γ or TAp63α isoforms, and identified all other molluscan p53 family sequences as p63 gene or its expression isoforms. Comparison of *M.galloprovincialis* p63 gene structure with other orthologues, found in sequenced genomes of four molluscs (*Aplysia californica*, *Lottia gigantea*, *Crassostrea gigas* and *Biomphalaria glabrata*), revealed the conserved p63 gene structure in the phylum Mollusca. Analyses also detected possible mis-annotations of the p63 gene in sequenced molluscan genomes and stressed the value of manual inspection (based on alignments of cDNA and protein onto the genome sequence) for a reliable and complete gene annotation.

(108 pages, 26 figures, 7 tables, 122 references, original in Croatian)

Keywords: Mollusca, Mytilus, orthologue, p53, p63, p73, genome, annotation

Supervisor: doc.dr.sc. Mauro Štifanić

Reviewers:
prof.dr.sc. Renato Batel
prof.dr.sc. Goran Klobučar
izv.prof.dr.sc. Bojan Hamer

Thesis accepted: 2nd June 2015

SADRŽAJ

1. UVOD 1

1. 1. Struktura i funkcija gena porodice p53 2

1.2. Geni porodice p53 kod beskralješnjaka 7

1.3. Osnovna obilježja mekušaca (koljeno Mollusca) i mediteranske dagnje *Mytilus galloprovincialis* 10

1.4. Ciljevi i polazne hipoteze istraživanja 13

2. LITERATURNI PREGLED 14

2.1. Uloge porodice p53 kod mekušaca (koljeno Mollusca) 15

2.2. Nukleotidni sljedovi porodice p53 mekušaca (koljeno Mollusca) 18

3. MATERIJALI I METODE 20

3.1. Uzorkovanje 21

3.2. Izolacija genomske DNA 21

3.3. Izolacija ukupne RNA i priprema cDNA biblioteke 22

3.4. Umnažanje DNA 24

3.5. Elektroforeza DNA 26

3.6. Izolacija iz gela i pročišćavanje PCR produkata 26

3.7. Ligacija u plazmidni vektor i transformacija bakterijskih stanica 26

3.8. Izolacija plazmidne DNA 27

3.9. Određivanje primarne strukture DNA (sekvenciranje) 28

3.10. Računalna analiza sljedova DNA i proteina 28

4. REZULTATI 29

4.1. Sljedovi kodirajuće DNA porodice p53 *Mytilus galloprovincialis* 30

4.2. Usporedba kodirajućih sljedova porodice p53 mekušaca 42

4.3. Genomski slijed gena porodice p53 *Mytilus galloprovincialis* 48

4.4. Očekivane ekspresijske izoforme gena porodice p53 *M. galloprovincialis* 64

4.5. Intronske regije 70

4.6. Usporedba strukture gena porodice p53 mekušaca 71

4.7. Filogenetska analiza 82

5. RASPRAVA 85

6. ZAKLJUČCI 93

7. LITERATURA 94

8. PRILOG 105

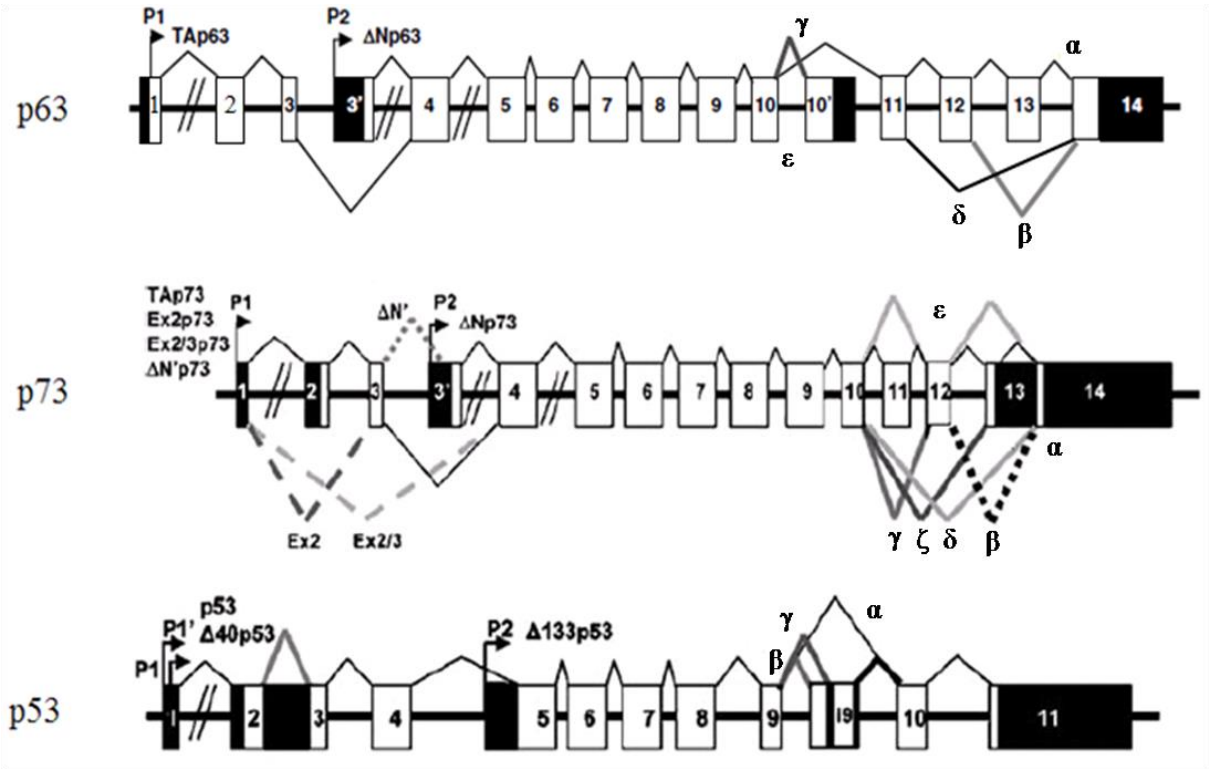
9. ŽIVOTOPIS VII

1. UVOD

1. 1. Struktura i funkcija gena porodice p53

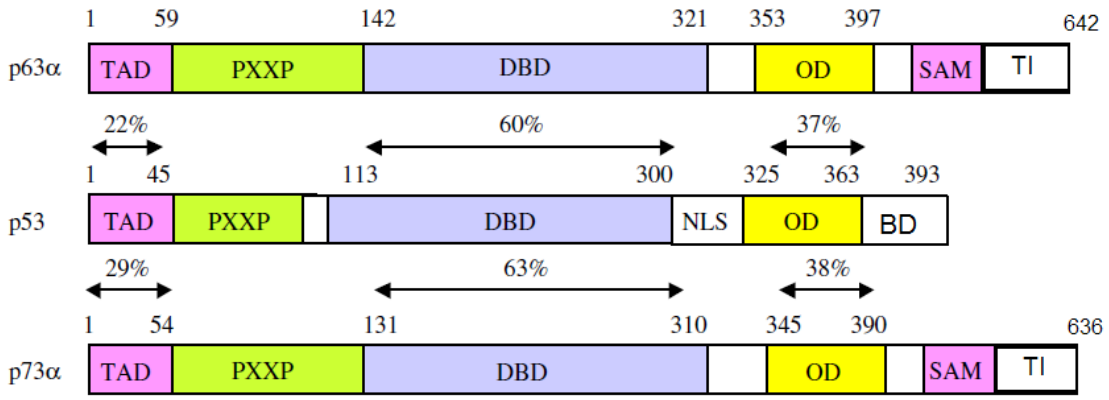
P53 porodica gena se kod kralješnjaka sastoji od tri člana; p53, p63 i p73. Geni porodice p53 su strukturni i funkcionalni homolozi s preklapajućim i zasebnim funkcijama (Strano i sur., 2001, Petitjean i sur., 2006). P63 i p73 su sličniji jedan drugome, nego genu p53 (Yang i sur., 1998, Melino i sur., 2003, Moll i Slade, 2004). U svojoj osnovi, proteini porodice p53 su transkripcijski faktori koji se specifično i izravno vežu kao tetrameri na ciljno mjesto u genomu i aktiviraju transkripciju gena smještenih u blizini mjesta vezanja (Bourdon i sur., 2005). Proteini kodirani tim ciljnim genima, a kojih je preko stotinu, na različite načine doprinose biološkim učincima porodice p53. Najznačajniji biološki učinci aktivnosti porodice p53 su regulacija staničnog ciklusa, programirane stanične smrti (apoptoze), staničnog starenja (senescencije), diferencijacije te ubrzan popravak DNA (Pietsch i sur., 2008, Qian i Chen, 2013). Kontroliranjem ovih važnih staničnih procesa, geni porodice p53 ostvaruju svoju ulogu tumor supresora tako što uklanjaju mogućnost nagomilavanja genetskih oštećenja koja doprinose tumorigenezi. U normalnim fiziološkim uvjetima, bez staničnog stresa, geni porodice p53 se nalaze u biološki latentnom stanju, eksprimirani u niskim koncentracijama i podvrgnuti brzoj ubikvitinaciji i proteolizi (Oren, 2003). Kad su stanice izložene različitim stresnim uvjetima, geni porodice p53 se aktiviraju. Aktivacija uključuje značajno povećanu ekspresiju proteina i postranslacijske modifikacije važne za brojne funkcije porodice p53. Najbolje istraženi stresni uvjeti u stanici koji aktiviraju gene porodice p53 su direktno DNA oštećenje i nepravilna stanična dioba iako su i drugi kao npr. nedostatak ribonukleotida i hipoksija također važni (D'Erchia i sur., 2003). Osim toga, brojni onkogeni proteini (Myc, Ras, adenovirus E1A i β-katenin) aktiviraju gene porodice p53 što direktno upućuje na njihovo tumor supresorsko djelovanje. Intenzivna temeljna i klinička istraživanja posljednjih 40-ak godina dokazala su da središnju ulogu u tumorigenezi i očuvanju genomskog integriteta ima gen p53 koji zbog toga i nosi naziv „čuvar genoma“. U genu p53 pronađene su mutacije u preko 50% svih humanih tumora (Hollstein i sur., 1994, Hollstein i sur., 1996). Geni p63 i p73 rijetko su mutirani u tumorima, a i tada su češće prekomjerno eksprimirani nego što dolazi do gubitka funkcije gena kao kod p53 (Bénard i sur., 2003). Ipak, p63 i p73 također sudjeluju u tumorigenezi, ali imaju važne i jedinstvene uloge i u procesima razvoja i diferencijacije (Irwin i Kaelin, 2001, Moll i sur., 2001, Danilova i sur., 2008).

Geni porodice p53 imaju vrlo sličnu egzon/intron organizaciju (Slika 1). Gen p53 čovjeka nalazi se na kromosomskom lokusu 17p13, veličine je 20 kb i obuhvaća 11 egzona odvojenih s 10 introna (Isobe i sur., 1986). Na 5' kraju gena p53 (kod svih sisavaca između egzona 1 i 2) prisutan je vrlo veliki intron (10 kb kod čovjeka). Prvi egzon gena p53 je uvijek nekodirajući, ali je uključen (formira se struktura petlje ili omče na mRNA) u regulaciju translacije (Mosner i sur., 1995). Gen p73 čovjeka nalazi se na kromosomskom lokusu 1p36, veličine je oko 70 kb i obuhvaća 14 egzona (Kaghad i sur., 1997a). Kod p73 čovjeka su prisutna tri velika introna : intron 1 (30 kb), intron 3 (25 kb) i intron 4 (14 kb). Egzon 1 gena p73 je, kao i kod p53, nekodirajući. Gen p63 čovjeka nalazi se na kromosomskom lokusu 3q28, veličine je oko 260 kb i obuhvaća 16 egzona (Yang i sur., 1998). P63 također ima dva velika introna (intron 1 i 3), ali prvi egzon je kodirajući. Egzon-intron struktura (organizacija) sva tri gena porodice p53 očuvana je kod različitih vrsta.



Slika 1. Shematski prikaz struktura gena porodice p53 (p53, p63 i p73) kod čovjeka. Egzoni su prikazani pravokutnicima i označeni brojevima (crna boja pravokutnika prikazuje nekodirajuće regije), a introni horizontalnim linijama Strjelice označavaju različita mjesta početka transkripcije. Na slici su prikazane i različite mogućnosti alternativnog izrezivanja introna. Preuzeto i prilagođeno iz Murray-Zmijewski i sur. (2006): *Cell Death Differ*. **13:** 962-972.

Slična organizacija gena porodice p53 odražava se u sličnoj strukturi proteina ove porodice. Sva tri člana (homologa, paraloga) porodice p53 sadrže tri očuvana područja (domene) (Slika 2). Na N-kraju nalazi se transaktivacijska domena (engl. *transactivation domain*, TA), središnje i najveće je DNA-vezna domena (engl. *DNA-binding domain*, DBD), a na C-kraju nalazi se oligomerizacijska domena (engl. *oligomerization domain*, OD).



Slika 2. Usporedna struktura proteina porodice p53. Prikazana su glavna funkcionalna područja: transaktivacijsko (TAD), područje bogato prolinom (PXXP), DNA-vezno područje (DBD), signalno područje za smještanje u jezgru (NLS), oligomerizacijsko područje (OD), područje sa sterilnim α-motivom (SAM), inhibitorno područje (TI) i bazno područje (BD). Označeni su postotci identičnosti pojedinih područja među članovima obitelji p53 kao i granični redni brojevi aminokiselina koje pojedinom području pripadaju. Preuzeto i prilagođeno iz Bénard i sur. (2003): *Hum. Mutat.* **21:** 182-191.

DNA-vezna domena pokazuje najveću homologiju među članovima porodice pa p63 i p73 u toj regiji imaju oko 60% identičnosti s p53. Osim toga, aminokiseline prepoznate kod p53 kao važne za strukturu (smatanje) DNA-vezne domene i interakciju sa DNA, očuvane su i kod p63 i p73. Zato se p63 i p73 uspješno vežu na poznata ciljna mjesta vezanja p53 u genomu (p53-REs; engl. *p53-responsive elements*) i aktiviraju prepisivanje (transkripciju) s tih promotora inducirajući zastoj u staničnom ciklusu, apoptozu ili stanično starenje (Levrero i sur., 2000, Melino i sur., 2003). Općeprihvaćen p53-RE konsenzus određen je kao dvije uzastopne kopije motiva „RRRCWWGYYY“ odvojene s razmakom od 0-13 pb, ali su kod nekih gena utvrđena i odstupanja od ovog motiva (Wang i sur., 2010). „R“ označava purine, „W“ baze A/T, a „Y“ pirimidine. Za razliku od DNA-vezne domene, područja na karboksi

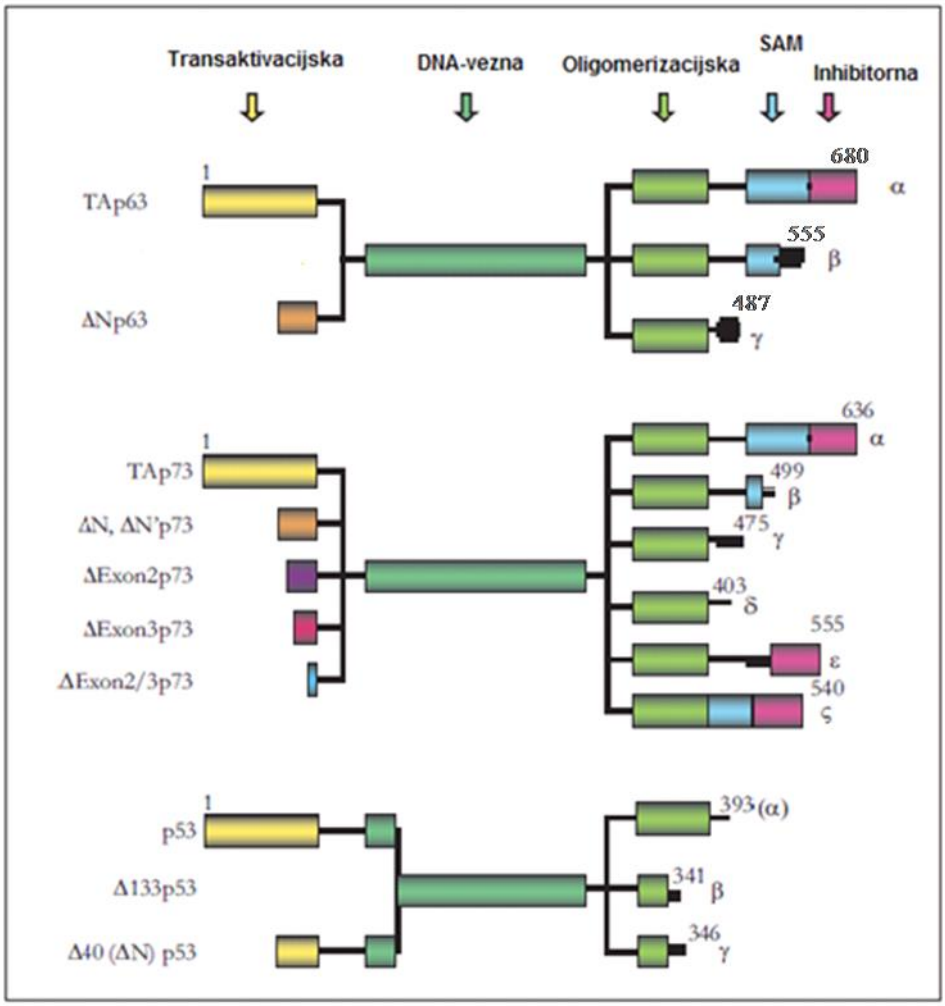
(C-) i amino (N-) krajevima p53 proteina imaju niz točno određenih mjesta fosforilacije i acetilacije, a koja nisu očuvana (konzervirana) kod p63 i p73. Tako je transaktivacijska domena najmanje konzervirana među članovima obitelji, sa 22% identičnosti između p63 i p53 te 29% identičnosti između p73 i p53. TA domena je mjesto vezanja koaktivatora transkripcije (npr. p300/CBP, TAFII40/60) i negativnih regulatora gena porodice p53 (npr. MDM2) te u interakciji s tim partnerima TA domena ostvaruje i stabilnu konformaciju alfa uzvojnice (Mavinahalli i sur., 2010). Za uspješno specifično vezanje na DNA i transkripcijsku aktivnost p53 protein mora biti u formi tetramera (dimer dimera), a tetramerizacija se ostvaruje preko oligomerizacijske domene. Proteini p63 i p73 također formiraju homotetramere slične 3D strukture kao p53 (Brandt i sur., 2009). Oligomerizacijska domena na C-kraju p53 ima 38% identičnosti sa p63 i p73 i ne formira heterotetramere s p63 i p73, ali oligomerizacijske domene p63 i p73 (identičnost 53%) međusobno tvore heterotetramere *in vitro* (Joerger i sur., 2009). Mogućnosti međusobnih interakcija članova p53 obitelji važne su za funkciju ovih proteina jer su u nekim tkivima oni eksprimirani istovremeno. Na svom C-kraju p63 i p73 imaju domenu sa sterilnim α-motivom (engl. *sterile α motif*, SAM) i transinhibitornu domenu (engl. *transactivation inhibitory domain*, TI), poznatu i kao post-SAM. P53 protein na svom C-kraju ima baznu domenu (engl. *basic domain*, BD). SAM domena je važna za protein-protein interakcije i karakteristična je za proteine uključene u prijenos signala u stanici i diferencijaciju (npr. Eph obitelj tirozin kinaznih receptora i ETS obitelj transkripcijskih faktora). SAM domene prepoznaju druge SAM domene te tvore homo-oligomere i hetero-oligomere ali vežu i druge proteine tako da SAM domene mogu poslužiti kao temelj izgradnje složenih proteinskih kompleksa u stanici (Kim i Bowie, 2003). Transinhibitorna (TI) se veže na transaktivacijsku (TA) domenu proteina i inhibira njegovu transkripcijsku aktivnost (Maisse i sur., 2003). Uloga bazne domene je najvjerojatnije pozitivna i negativna regulacija p53 aktivnosti jer je gotovo svaka aminokiselina bazne domene često mjesto posttranslacijskh modifikacija kao što su fosforilacija, metilacija, acetilacija, ubikvitinacija i sl. (Scoumanne i sur., 2005). Između DNA-vezne i oligomerizacijske domene se kod p53 nalazi malo područje-signal za smještanje u jezgru (engl. *nuclear localization signal*, NLS). Svi predstavnici porodice p53 sadrže i bar jedno područje bogato prolinom (PRD; engl. *proline rich domain*) s motivom PXXP (gdje je P=prolin, a X=bilo koja druga aminokiselina).

Svi članovi porodice p53 imaju dva promotora: P1 smješten uzvodno od egzona 1 i P2 smješten unutar introna 3 (p63 i p73) odnosno introna 4 (p53). Korištenjem dvaju promotora i alternativnim izrezivanjem, proteini p53, p63 i p73 se mogu eksprimirati u više funkcionalno različitih izoformi (Slika 1). Sve izoforme imaju DNA-veznu domenu, ali se razlikuju u regijama N- i C-kraja proteina (Slika 3). Prepisivanjem gena s promotora P1 nastaju izoforme koje imaju transaktivacijsku domenu (p53, TAp63 i TAp73). Prepisivanjem gena s promotora P2 nastaju ΔN izoforme s posebnim aminokiselinskim slijedom (egzon 3' kod p63 i p73) na N-kraju (ΔNp63 i ΔNp73) ili bez tog slijeda aminokiselina (Δ133p53). Ovoj Δ133p53 izoformi p53 nedostaje čak i dio DNA-vezne domene. Uz korištenje različitih promotora (P1 i P2) na 5' kraju gena porodice p53 događa se i alternativno izrezivanje primarnih transkripata gena. Na taj način nastaju tri izoforme p73 kojima nedostaje egzon 2, egzon 3 ili oba ova egzona (Ex2p73, Ex3p73 i Ex2/3p73) te jedna izoforma koja nema izrezan intron 3 u svom transkriptu (ΔN'p73). Kod gena p53 alternativnim izrezivanjem introna 2 ili alternativnom inicijacijom translacije nastaje Δ40p53 izoforma kojoj nedostaje dio transaktivacijske domene. Izoforme porodice p53 kojima nedostaje transaktivacijska domena (ΔN izoforme) najčešće djeluju kao dominantno negativni inhibitori izoformi s cjelovitom TA domenom (p53, TAp63, Tap73). Tako ΔNp73 inhibira p53 kompeticijom za DNA vezna mjesta, a p73 direktnim protein-protein interakcijama (Tozluoğlu i sur., 2008).

Većina alternativnog izrezivanja događa se ipak na 3' kraju, od egzona 10 do 13, čime nastaju izoforme koje se razlikuju na svom C-kraju. Za p63 poznato je pet izoformi; α, β, γ, δ i ε, a za p73 šest različitih izoformi C-kraja; α, β, γ, δ, ε i ζ (Murray-Zmijewski i sur., 2006, Mangiulli i sur., 2009). γ izoforme p63 i p73 najsličnije su proteinu p53 jer im nedostaje većina C-kraja. Samo α izoforme sadrže SAM domenu. p63δ nastaje izrezivanjem egzona 12 i 13, a p63ε prijevremenom terminacijom transkripcije u intronu 10. Poznate su i izoforme C-kraja p53 (Bourdon i sur., 2005). Alternativnim izrezivanjem introna 9 gena p53 nastaju tri različite izoforme, po analogiji nazvane α, β i γ od kojih β i γ nemaju cjelovito oligomerizacijsko područje.

Kompleksna struktura porodice p53 i regulatornih mehanizama razlog su složenog međudjelovanja članova porodice p53 i njihovih izoformi u brojnim biološkim funkcijama. Cjelovite (TA) i skraćene (ΔN) izoforme porodice p53 uglavnom imaju suprotne biološke funkcije: skraćene izoforme potiču diobu stanica, a TA izoforme zastoj u staničnom ciklusu, stanično starenje i apoptozu. Ipak, aktivnost proteina porodice p53 također ovisi i o uvjetima

unutar stanice te često ravnoteža između različitih izoformi određuje sudbinu stanice (DeYoung i Ellisen, 2007, Collavin i sur., 2010).



Slika 3. Shematski prikaz glavnih proteinskih izoformi porodice p53. Različitim bojama prikazane su glavne funkcionalne domene; transaktivacijska, DNA-vezna, oligomerizacijska i domena sa sterilnim α-motivom (SAM). Uz svaku izoformu naznačen je broj aminokiselina (cjelovite izoforme; TA). Preuzeto i prilagođeno iz Petitjean, A. i sur. (2006): *Bull. Cancer* **93(12):** E126-35.

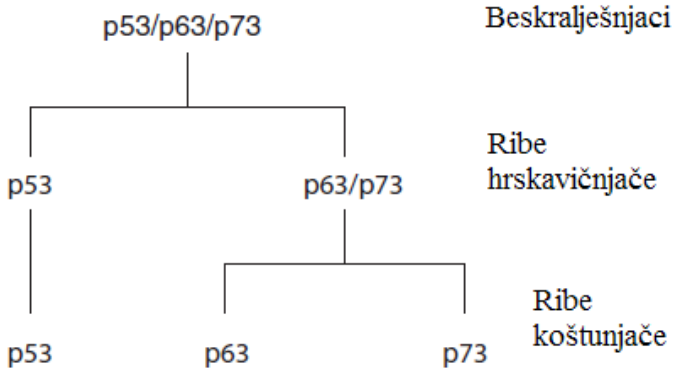
1.2. Geni porodice p53 kod beskralješnjaka

Gene porodice p53 ne nalazimo samo kod kralješnjaka, nego i kod mnogih vrsta beskralješnjaka. Ipak, kod beskralješnjaka još uvijek često nedostaju podaci o strukturi,

molekularnim mehanizmima i funkciji homologa gena p53. Najčešće su poznati tek cDNA sljedovi i eksprimirane proteinske izoforme porodice p53 prisutne u različitim (stresnim) uvjetima kojima su organizmi bili izloženi. Cjeloviti sljedovi gena porodice p53 kod nekih beskralješnjaka potvrđeni su zahvaljujući brojnim projektima sekvenciranja cjelokupnih genoma. Za sada je porodica p53 kod većine genoma beskralješnjaka zastupljena samo jednim ili sa više gena homologa (Nedelcu i Tan, 2007). Takvi ishodišni geni uvijek imaju u izvedenim proteinskim sljedovima očuvanu DNA-veznu domenu, a prema tome imaju li ili ne SAM domenu klasificiraju se kao geni slični p53 (bez SAM domene) ili p63/73 (sa SAM domenom). Tako morski ježinac *Strongylocentrotus purpuratus* ima u svom genomu jedan gen porodice p53 i on je sličan p63/73, kao i mnogočetinaš *Capitella teleta* (Simakov i sur., 2013). U sekvenciranim genomima kestenjastog brašnara *Tribolium castaneum* i obične vodenbuhe *Daphnia pulex*, također je porodica p53 zastupljena jednim genom. Ipak, kod najjednostavnijeg predstavnika Metazoa, jednostaničnog bičaša *Monosiga brevicollis* prisutna su dva gena i to jedan sličan p53, a drugi p63/73 (King i sur., 2008). Više gena homologa porodice p53 pronađeno je i kod žarnjaka *Nematostella vectensis* (tri gena slična p53, (Pankow i Bamberger, 2007)), plošnjaka *Schistosoma mansoni* (dva gena slična p53, (Berriman i sur., 2009)), kukca *Anopheles gambiae* (dva gena slična p53, (Holt i sur., 2002)), plaštenjaka *Ciona intestinalis* (dva gena slična p53, (Putnam i sur., 2007)) i svitkoglavca *Branchiostoma floridae* (jedan sličan p53, a drugi p63/73 (Putnam i sur., 2008)). Određivanje koliko je gena homologa prisutno u genomu ima mnoga ograničenja kao što su kvaliteta sekvenciranog genoma i različit stupanj očuvanosti pojedinih domena porodice p53. Osim toga, broj gena porodice p53 nekog organizma može biti rezultat naknadnih duplikacija ili delecija. Za pouzdanu interpretaciju evolucije gena porodice p53 potrebna su detaljnija istraživanja strukture i funkcije gena porodice p53 kod beskralješnjaka.

Struktura i uloga gena porodice p53 kod beskralješnjaka opširnije je istražena tek kod dva važna modelna organizma; oblića *Caenorhabditis elegans* (Derry i sur., 2001, Schumacher i sur., 2001) i vinske mušice *Drosophila melanogaster* (Ollmann i sur., 2000). Genom *C. elegans* sadrži jedan gen porodice p53, nazvan *cep-1*, a *D. melanogaster* također jedan gen, nazvan *dmp53*. Protein Cep-1 na svom C-kraju ima SAM domenu, a Dmp53 nema. Analiza funkcije proteina CEP-1 i Dmp53 pokazala je da su oni potrebni za aktivaciju apoptoze spolnih stanica kao odgovor na oštećenje DNA (Brodsky i sur., 2004, Schumacher i sur., 2004, Derry i sur., 2006). Na taj način, čak i beskralješnjaci kao *C.elegans*, koji nemaju

obnavljajućih tkiva i imaju životni vijek od 2-3 tjedna te tako nisu podložni razvoju tumora, moraju nadzirati gensku stabilnost svojih spolnih stanica i ukloniti one sa oštećenom DNA. Jedan od proteina porodice p53 žarnjaka *Nematostella vectensis* također je potreban za aktivaciju apoptoze spolnih stanica nakon izloženosti jedinki UV zračenju (Pankow i Bamberger, 2007). Očuvanje genetske stabilnosti spolnih stanica smatra se ishodišnom funkcijom proteina porodice p53. U tom smjeru, zanimljivo je otkriće da je TAp63α odgovoran za održavanje genske stabilnosti oocita sisavaca i uklanjanje oštećenih stanica (Suh i sur., 2006) što upućuje na to da je p63 evolucijski ishodišni gen porodice p53. Kasnije u evoluciji pojavili su se p73 i p53 kao tumor supresori u svim tjelesnim stanicama. Smatra se da se po prvi puta jasno odvojen gen p53 pojavio kod riba hrskavičnjača, a kod riba koštunjača dogodilo se odvajanje gena p63 i p73 (Slika 4). Nakon toga, dogodilo se vrlo malo evolucijskih promjena u genima p63 i p73. Za razliku od njih, evolucija gena p53 je intenzivna i brza, pribavljajući nove funkcije u nadzoru genomske stabilnosti svih tjelesnih stanica. Ovaj korak od vrlo specifičnog mehanizma kontrole spolnih stanica do općenitog nadzora i svih ostalih tjelesnih stanica postao je neophodan jer se životni vijek jedinke znatno produžio te premašuje život same stanice i obnova tkiva postaje nužna (Belyi i sur., 2010, Dotsch i sur., 2010, Rutkowski i sur., 2010).



Slika 4. Pretpostavka evolucije gena porodice p53. Najraniji homolog porodice p53 javlja se već kod beskralješnjaka, a jasno odvajanje u tri zasebna gena dogodilo se kod riba koštunjača. Preuzeto i prilagođeno iz Belyi, V.A. i sur. (2010): *Cold Spring Harb. Perspect. Biol.* **2:** a001198.

1.3. Osnovna obilježja mekušaca (koljeno Mollusca) i mediteranske dagnje *Mytilus galloprovincialis*

Mekušci su vrlo brojna i raznolika skupina beskralješnjaka. Mogli bi ih opisati kao životinje s nekolutićavim, bilateralno simetričnim ili sekundarno asimetričnim tijelom. Naziv Mollusca (mekušci) su dobili prema latinskoj riječi „molis“ što znači mekan, gibak, nježan, a odnosi se na mekano tijelo što se nalazi unutar čvrste ljuske. Mekano tijelo obavija kožni nabor- plašt, koji u većine mekušaca izlučuje ljušturu iz organske tvari konhiolina i kalcij-karbonata. Ljuštura je izgrađena od osam, dva ili jednog dijela. Na tijelu se razlikuju glava, stopalo i trup. Imaju srce i krvne žile, škrge za disanje, a u nekih dio plaštene šupljine obavlja ulogu pluća. Spolovi su im razdvojeni, ali ima i dvospolaca. Opisano je preko 50000 živućih vrsta, a brojne su i fosilne vrste. Koljenu Mollusca pripadaju neki od najstarijih poznatih mnogostaničnih životinja (Metazoa). Žive u moru i u kopnenim vodama te na kopnu. Ljudi mekušce upotrebljavaju za hranu, a od njihovih ljušura i kućica često izrađuju ukrase. Mekušcima pripada sedam razreda: trbožlijepci (Aplacophora), jednoljušturaši (Monoplacophora), babuške (Polyplacophora), školjkaši (Bivalvia), koponošci (Scaphopoda), puževi (Gastropoda) i glavonošci (Cephalopoda) (Matoničkin, 1998).

Mediteranska dagnja *Mytilus galloprovincialis* (Lamarck, 1819) je morski mekušac (koljeno Mollusca) iz razreda školjkaša (Bivalvia) (Slika 5). U Europi je rasprostranjena u Sredozemnom i Crnom moru te na istočnoj obali Atlantskog oceana od juga Španjolske pa sjeverno do obale Velike Britanije (Gosling, 1992). Jedna je od najvažnijih akvakulturnih vrsta u Europi. Još su dvije srodne vrste roda *Mytilus*: *M. edulis* (obala Atlantskog oceana od Francuske do Norveške) i *M. trossulus* (Baltičko more i obale Škotske). Izgledom je mediteranska dagnja velika školjka (do 12 cm) plavo-crne boje ljuštura. Sesilni (sjedilački) je organizam, a najčešće nastanjuje zonu plime i oseke stjenovitih obala te raste u gustoj zajednici s mnogo jedinki. Za podlogu se pričvršćuje bisusnim nitima koje izlučuje mišićavo stopalo dagnje.

Slika 5. Istraživani organizam u ovom radu, mediteranska dagnja *Mytilus galloprovincialis* na svom prirodnom staništu u moru.

Mediteranska dagnja je dobro prilagođena promjenama različitih ekoloških parametara (salinitet, temperatura, količina kisika, dostupnost hrane) te naseljava i čista i zagađena područja. Hrani se suspendiranim česticama (fitoplankton, fini organski detritus, razgrađena organska tvar, bakterije, mikrozooplankton i anorganske čestice) neselektivno filtrirajući velike količine mora te na taj način uz hranu nakuplja (akumulira) i različite kemijske tvari (zagađivala) u svojim tkivima. Zbog svojih ekoloških obilježja (način života i prehrane) i važne uloge u hranidbenom lancu mediteranska dagnja (ali i drugi školjkaši) koristi se kao bioindikatorski organizam u studijama praćenja stanja kvalitete morskog okoliša (monitoring). Uobičajena mjerenja koja se koriste u monitoringu okoliša uključuju opterećenost organizama kemikalijama (zagađivalima) i fiziološke markere morfologije, reprodukcije, obrambenih mehanizama i oštećenja tkiva . Prate se patološke promjene za procjenu zdravlja i vitalnosti organizama osobito u onečišćenim područjima (Vukmirovic i sur., 1994, Hamer i sur., 2004, Dondero i sur., 2006, Gagné i sur., 2008, Hamer i sur., 2008, Banni i sur., 2009, Štambuk i sur., 2013, Jarque i sur., 2014). Škrge su kao organ kontinuirano i direktno izložene djelovanju zagađivala iz mora pa se one najčešće odabrane kao tkivo za daljnje analize. Kako bi se unaprijedio biomonitoring sve je češća upotreba genetike u

istraživanjima jer pruža brze i pouzdane metode praćenja utjecaja onečišćenja na prirodne populacije.

Za dagnju *Mytilus galloprovincialis* veličina genoma (C-vrijednost) tj. ukupna količina DNA sadržana u jezgri haploidne stanice (n=14) procijenjena je na 1,41-1,92 pg (http://www.genomesize.com) što odgovara veličini od 1,4-1,9 x 10^9 pb. Mitohondrijski genom obuhvaća još 17,1 kb i kodira za 37 gena. Poznati sljedovi gena dostupni u bazama podataka dosad su najčešće rezultat istraživanja pojedinačnih gena vezanih uz bolesti, otpornost i imunost jer su takvi geni od interesa u akvakulturi. Kako se dagnje intenzivno koriste u biomonitoringu, poznati su i geni uključeni u odgovor na zagađenje i stres. Mnogo je podataka u bazi vezano i za ribosomalne gene i pojedine mitohondrijske gene zbog njihove upotrebe u istraživanjima populacijske genetike dagnje i filogenije školjkaša (Saavedra i Bachère, 2006). Međutim, još je nedovoljno poznat genom i svi geni dagnje *M. galloprovincialis* što često ograničava i znanje o molekularnim osnovama nekih važnih procesa kao što su regulacija rasta i diferencijacije ili spolno sazrijevanje dagnje. Za analize ekspresije gena korisne su metoda mikročipa (engl. *microarray*) i baze sekvenciranih fragmenata ekprimiranih nukleotidnih sljedova (engl. *expressed sequence tag*, EST) (Venier i sur., 2003, Tanguy i sur., 2008).

Zadnjih godina značajan je napredak u broju i raznolikosti sekvenciranih genoma. U NCBI GenBank bazi podataka je sekvencirano i javno dostupno četiri genoma mekušaca: puževa *Aplysia californica* (http://www.ncbi.nlm.nih.gov/bioproject/PRJNA209509), *Lottia gigantea* (http://genome.jgi-psf.org/cgi-bin/dispGeneModel?db=Lotgi1&id=182533) i *Biomphalaria glabrata* (http://genome.wustl.edu/genomes/detail/biomphalaria-glabrata/) te školjkaša *Crassostrea gigas* (http://www.ncbi.nlm.nih.gov/bioproject/PRJNA70283). Informacije o nukleotidnim sljedovima sva četiri sekvencirana genoma mekušaca još nisu pridružene pojedinim kromosomima. Genom *Aplysia californica* dodatno je obrađen kao NCBI RefSeq genom. Analiza genoma mekušaca omogućava određivanje strukture gena mekušaca, a time i pouzdanu pretpostavku mogućih proteinskih izoformi, funkcija gena i sl. Osim toga, sekvencirani genomi pružaju mogućnost istraživanja cjelokupnog mehanizma djelovanja gena mekušaca i pronalaženja svih sudionika (aktivatora, inhibitora, i sl.) u tom složenom mehanizmu.

1.4. Ciljevi i polazne hipoteze istraživanja

Ciljevi:

1. Utvrditi koji su članovi porodice p53 prisutni kod dagnje *Mytilus galloprovincialis*
2. Doznati potpuni nukleotidni slijed gena porodice p53 dagnje *Mytilus galloprovincialis*
3. Opisati strukturu gena uključujući raspored i broj egzona i introna te regulatorne regije i motive
4. Usporediti opisane sljedove s poznatim sljedovima gena porodice p53 iz drugih organizama

Hipoteze:

1. Dagnja ima jedan-ishodišni gen porodice p53
2. Po svojoj strukturi gen porodice p53 kod dagnje najsličniji je genu p63 kod kralješnjaka
3. Domene i regulatorne regije, opisane kod kralješnjaka, očuvane su i kod dagnje *Mytilus galloprovincialis*

2. LITERATURNI PREGLED

2.1. Uloge porodice p53 kod mekušaca (koljeno Mollusca)

Porodica p53 nametnula se kao dobar kandidat za korištenje u biomonitoringu okoliša pa su brojna istraživanja strukture i funkcije ove obitelji gena u različitim vrstama koljena Mollusca pokrenuta upravo s ciljem razvoja gena porodice p53 kao biomarkera genotoksičnog onečišćenja. Iako su homolozi porodice p53 utvrđeni kod različitih vrsta koljena Mollusca, malo se zna o aktivnosti i funkcijama tih homologa. Podaci o ekspresiji poznati su za nekoliko morskih školjkaša i prikazuju odnose između razina mRNA i promjena u okolišnim parametrima ili eksperimentalnim uvjetima. Uloga porodice p53 kod mekušaca opsežnije je istraživana u neoplaziji hemocita. Neoplazija hemocita je proliferativna bolest koja se javlja u nekih morskih školjkaša. Ova bolest je složen poremećaj krvnih stanica školjkaša (hemocita) koji karakterizira brzo umnažanje stanica koje se ne prihvaćaju za podlogu, neprekidno se dijele, imaju udio DNA u stanici 3.8-7.7 puta veći od diploidnog te jezgra dominira u odnosu na citoplazmu u stanici (Barber, 2004, Ruiz i López, 2013). Takvi neoplastični hemociti prodiru u ostala tkiva (gonada, plašta, stopala i sl.) te se bolest klasificira kao leukemija. Uzrok razvoja neoplazije hemocita se još ne može sa sigurnošću odrediti, ali je potvrđeno da su školjke izložene okolišnom onečišćenju podložnije razvoju ovog poremećaja (St.-Jean i sur., 2005). Razvoj genetskih biomarkera za neoplaziju hemocita omogućio bi brzo i precizno pregledavanje populacija školjkaša i praćenje stanja okoliša. Tako se pokušalo utvrditi specifične antigene (markere) normalnih i stanica leukemije prema proteinskim profilima normalnih i leukemija stanica kod školjkaša *Mya arenaria* (Stephens i sur., 2001). Uočene su značajne razlike u proteinskim profilima ova dva tipa stanica. Stanice leukemije u završnoj fazi eksprimiraju jedinsteveni antigen IE10 (vrlo hidrofobni 252 kDa glikoprotein, vjerojatno transmembranski protein), a normalne stanice eksprimiraju 2A4 marker (184 kDa glikoprotein). Homolozi porodice p53 su također različito eksprimirani u normalnim i stanicama leukemije. Koristeći antitijelo specifično za centralnu DNA vezujuću domenu proteina, u hemocitima vrste *M. arenaria* uočeno je da je homolog nazvan Map73 (74,8 kDa) eksprimiran samo u stanicama leukemije dok je homolog Map53 (52,8 kDa) eksprimiran i u normalnim i u stanicama leukemije. Osim toga, imunokemijska analiza je pokazala da su u normalnim hemocitima homolozi p53 obitelji lokalizirani uglavnom u jezgri, a u hemocitima leukemije isključivo u citoplazmi (Kelley i sur., 2001, Stephens i sur., 2001). Neoplazija

hemocita školjkaša može poslužiti kao koristan model za istraživanje funkcija gena porodice p53 jer pruža *in vitro* i *in vivo* alternativu humanim leukemija staničnim linijama.

Metodom kvantitativne lančane reakcije polimerazom u stvarnom vremenu (engl. *Quantitative Real-Time PCR*; Q-PCR) kod školjkaša *Spisula solidissima* pratila se ekspresija tri p53 homologa (nazvana p53, p97 i p120) tijekom embrionalnog razvoja jedinke i promjene u ekspresiji nakon izloženosti polikloriranim bifenilima (eng. *Polychlorinated biphenyls*, PCBs). Uočena je prisutnost p53 i p97 u embrionalnom i odraslom tkivu dok je p120 isključivo prisutan u embrionalnom tkivu i najviša razina ekspresije p120 mRNA podudara se sa početkom neuralnog razvoja školjkaša *Spisula solidissima*. Kako se embrij razvija, p53 i p97 su konstitutivno eksprimirani dok se razina p120 mRNA mijenja. Najviša razina ekspresije mRNA p120 je nakon 24h, a nakon toga počinje padati. Suprotno, do 24h protein p120 uopće nije detektiran, ali nakon 48h njegova razina raste što upućuje na to da je p120 vjerojatno reguliran na razini translacije. Homolozi p53 školjkaša *S. solidissima* nisu samo različito eksprimirani tijekom normalnog razvoja, nego pokazuju različite odgovore i nakon izloženosti okolišnim stresorima. Nakon izloženosti embrija PCB-u, najznačajnije je suprimirana ekspresija p120 dok p97 pokazuje manje i kasnije smanjenje ekspresije, a razina p53 nije se mijenjala (Jessen-Eller i sur., 2002). Istom metodom analizirana je i tkivna ekspresija (probavna žlijezda, škrge, plašt, mišić, hemolimfa) p53 homologa dagnje *Mytilus galloprovincialis*. P53 homolog je eksprimiran u svim ispitivanim tkivima. Relativna količina ekspresije izražena je kao omjer sa razinom ili 18S rRNA ili aktina te pokazuje najvišu razinu ekspresije u škrgama pa onda u probavnoj žlijezdi, plaštu, mišiću i hemocitima. Razlike nisu vrlo značajne, unutar jednog reda veličine. Tada je u istim tkivima promatran utjecaj subletalne doze poznatog karcinogena benzo(a)pirena (B(a)P). Nakon 24h izloženosti dozi od 75nM uočava se povećana razina ekspresije u probavnoj žlijezdi i značajno smanjenje u hemocitima dok nisu zapažene promjene u škrgama, plaštu i mišiću (Banni i sur., 2009). Nakon izloženosti (6 i 8 dana) dagnji roda *Mytilus* antropogenim zagađivalima žive (Hg) i sirove nafte ("*North Sea oil*", NSO) uočeno je, metodom mikročipa, a potvrđeno i kvantitativnim PCR-om u stvarnom vremenu, smanjenje ekspresije p53 homologa u tkivu probavne žlijezde (Dondero i sur., 2006).

Kod dagnje *M. galloprovincialis* istraživane su metodom mikročipa i sezonske promjene u ekspresiji u tkivima probavne žlijezde i gonada ženskih i muških jedinki uzorkovanih na njihovom prirodnom staništu. Tada je uočen porast ekspresije p53 homologa dagnje u ljetnim

mjesecima (kolovoz). Uz p53, porast je ekspresije i drugih gena vezanih uz imunološki odgovor (npr. matrilin) pa je ovaj rezultat objašnjen kao vjerojatna prilagodba na moguće infekcije toksičnim dinoflagelatima koje su najčešće u tom razdoblju godišnjeg ciklusa (Banni i sur., 2011). Kod školjkaša *Crassostrea gigas* p53 homolog je eksprimiran u svim tkivima (prvi pokus Q-RT-PCR). Razina p53 mRNA je oko dva puta veća u mišiću aduktoru, plaštu i hemolimfi, nego u škrgama, probavnoj žlijezdi i labijalnim palpima. Uočene su i sezonske promjene razine ekspresije u tkivu škrga, ali i ukupnom mekom tkivu školjki sa najvišim vrijednostima u zimskom periodu; Siječanj-veljača-ožujak 2004. i veljača-ožujak-travanj 2005, kada je temperatura mora bila i ispod 10°C. Općenito najniže vrijednosti mRNA ekspresije zabilježene su u rujnu u oba tipa tkiva. Zanimljivo je da je tijekom mjerenja uočena i varijabilnost ekspresije među jedinkama. Čimbenik koji barem djelomično to objašnjava bio bi spol jedinke (Farcy i sur., 2008).

Iz brojnih istraživanja se uvidjela strukturalna sličnost porodice p53 kod mekušaca i čovjeka pa se eksperimentalno željelo potvrditi da p53 homolog školjkaša *Mya arenaria* i p53 čovjeka i djeluju na isti način u stanici. Plazmidi koji eksprimiraju ili ljudski p53 (Hsp53) ili p53 homolog školjkaša *M. arenaria* (Map53) unešeni su trasfekcijom u p53-null H1299 staničnu liniju dobivenu od ljudske jetre. Funkcionalnost proteina izražena je praćenjem p53/Mdm2 interakcije i ekspresije p53-reguliranih gena staničnog ciklusa (npr. CIP1). Dok je ekspresija humanog p53 spontano inducirala ekspresiju CIP1, ekspresija p53 školjkaša to nije (Holbrook i sur., 2009).

Primjer p53 homologa kod havajske sipice *Euprymna scolopes* otkriva da osim dobro poznate uloge u odgovoru na okolišno zagađenje, članovi porodice p53 su i važni čimbenici u razvojnim procesima jedinke. Nekoliko sati nakon što se mlada jedinka izleže iz jaja, sipa *E. scolopes* formira specifičan svjetlosni organ kroz simbiozu sa morskom bioluminiscentnom bakterijom *Vibrio fischeri.* Interakcija sa simbiontom rezultira gubitkom složenog trepetljikavog epitela procesom apoptoze. Istraživanja su dokazala (QRT-PCR, imunocitokemija i western blot) prisutnost p53 homologa u stanicama koje formiraju svjetlosni organ te se prepostavlja da je “stres” početka simbioze signal koji aktivira p53 homologa koji inicira apoptozu i razvojnu morfogenezu svjetlosnog organa. Nakon početka simbioze dokazana je veća razina proteina u jezgama stanica svjetlosnog organa nego u citoplazmi kao rezultat nakupljanja p53 homologa (nema povećanja transkripcije) što upućuje na aktivaciju p53 homologa u ovim stanicama (Goodson i sur., 2006). Sva ova istraživanja

zapravo ne utvrđuju koja je stvarna uloga i svrha p53 homologa u uočenim promjenama, nego samo najčešće detektiraju promjenu u ekspresiji p53 homologa. Za daljnja istraživanja funkcije potrebno je razriješiti strukturu gena homologa porodice p53.

2.2. Nukleotidni sljedovi porodice p53 mekušaca (koljeno Mollusca)

Istraživanja funkcije p53 homologa kod mekušaca rezultirala su pohranom većeg broja nukleotidnih sljedova p53 homologa u GenBank bazu (http://www.ncbi.nlm.nih.gov/genbank/). Pretraživanjem GenBank baze nukleotidnih sljedova prema ključnim riječima "Mollusca" i "p53" i prema međusobnoj sličnosti sekvenci, mogu se pronaći svi javno dostupni poznati nukleotidni sljedovi p53 homologa kod različitih vrsta mekušaca. Dva su genomska slijeda poznata; cjeloviti školjkaša *Mya arenaria* nazvan "p63/p73 and p53 gene" (FJ041332, nije publicirano), dužine 16798 nt i djelimičan dagnje *Mytilus galloprovincialis* nazvan "gene for putative p53 tumor suppressor, egzons 1-2" (AJ9666664, nije publicirano), dužine 1333 nt. Svi ostali poznati sljedovi porodice p53 mekušaca su sljedovi kodirajuće DNA. Prvi identificirani p53 homolog koljena Mollusca je u pučinske lignje *Loligo forbesi* (nije publicirano). Dva su poznata kodirajuća DNA slijeda *Loligo forbesi*, oba nazvana Sqp53; cjeloviti (U43595) dužine 1828 nt i nepotpuni (U43596) dužine 1316 nt. Kod havajske lignje *Euprymna scolopes* poznata su dva cjelovita slijeda (Goodson i sur., 2006) nazvana p63(47) (DQ247973) i p63(62) (DQ247974). p63(47) dug je 1296 nt, a p63(62) 1701 nt. Za školjkaša *Spisula solidissima* poznata su tri slijeda nazvana redom: p120 (AF285104, (Jessen-Eller i sur., 2002)), Ssp63/73α (AY289767, (Cox i sur., 2003)) i Ssp63/73β (AY289768, (Cox i sur., 2003)). Ssp63/73α je 2699 nt dug slijed, a Ssp63/73β 3920 nt. Nukleotidni sljedovi ove dvije varijante (homologa) su skoro identične do stop kodona, ali razlikuju se u 3' netranslatiranoj regiji (eng. *Untranslated region*, UTR). Zaključeni aminokiselinski slijed obje varijante dug je 597 aminokiselina i odgovara proteinu molekularne mase oko 68 kDa.

Za školjkaša *Mya arenaria* poznata su dva cjelovita slijeda (Kelley i sur., 2001): "p53-like" (AF253323), dužine 2541 nt i "p73-like" (AF253324) dužine 2455 nt. Poznat je jedan cjeloviti slijed školjkaša *Crassostrea gigas*, nazvan "p53-like" (AM236465, (Farcy i sur., 2008)). Ova cjelovita kodirajuća DNA je 2116 nt duga od čega je 1305 nt otvoreni okvir čitanja kojim je kodiran protein od 434 aminokiseline i molekularne mase 49 kDa. Jedan

cjeloviti kodirajući slijed, nazvan “p53-like”, poznat je i za puža *Biomphalaria glabrata* (GU929337), a dužine je 1283 nt (Baeza Garcia i sur., 2010). Za puža Haliotis tuberculata poznat je također jedan cjeloviti kodirajući slijed, “p73-like” (AM396936, nije publicirano) dužine 2192 nt i otvorenim okvirom čitanja od 614 aminokiselina. Najviše poznatih kodirajućih DNA sljedova p53 homologa koljena Mollusca je od roda *Mytilus*. Zastupljene su vrste *Mytilus edulis*, *Mytilus trossulus* i *Mytilus galloprovincialis* sa ukupno 13 kodirajućih DNA sljedova. Tri su djelimična slijeda; “p53-like” *M. galloprovincialis* (DQ158079, (Dondero i sur., 2006)), “p53-like”3’UTR *M. galloprovincialis* (EF080937, nije publicirano) i p53 *M. edulis* (AY705932, (Ciocan i Rotchell, 2005). Ostalih deset je cjelovitih sljedova i to *M. edulis* i *M. trossulus* (Muttray i sur., 2005, Muttray i sur., 2007) te se prema nazivima mogu grupirati u sljedove “p53-like” (AY579472 i AY611471), “p63/73-like” (DQ865150, DQ060435, DQ865152 i DQ060437) i “ΔNp63/73-like” (DQ865151, DQ060436, DQ865153 i DQ060438). Zbog sličnosti gena porodice p53 i velikog broja mogućih izoformi, nomenklatura sljedova porodice p53 mekušaca je, kao i kod većine beskralješnjaka, vrlo zamršena i nejasno definirana.

3. MATERIJALI I METODE

3.1. Uzorkovanje

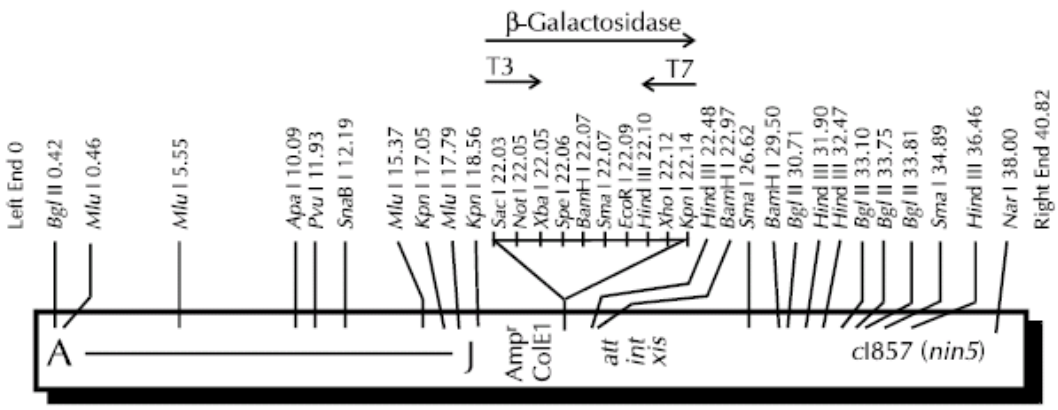
Jedinke dagnje *Mytilus galloprovincialis* su sakupljene (uzorkovane) direktno na svom prirodnom staništu u moru, na lokalitetu uvale Valdibora (45°5' N, 13°38'E) u Rovinju u ožujku 2008. i 2011. godine. Prosječna veličina uzorkovanih dagnji bila je D/Š/V = 40/20/20. Iz svake jedinke izolirani su uzorci tkiva škrga za daljnje analize.

3.2. Izolacija genomske DNA

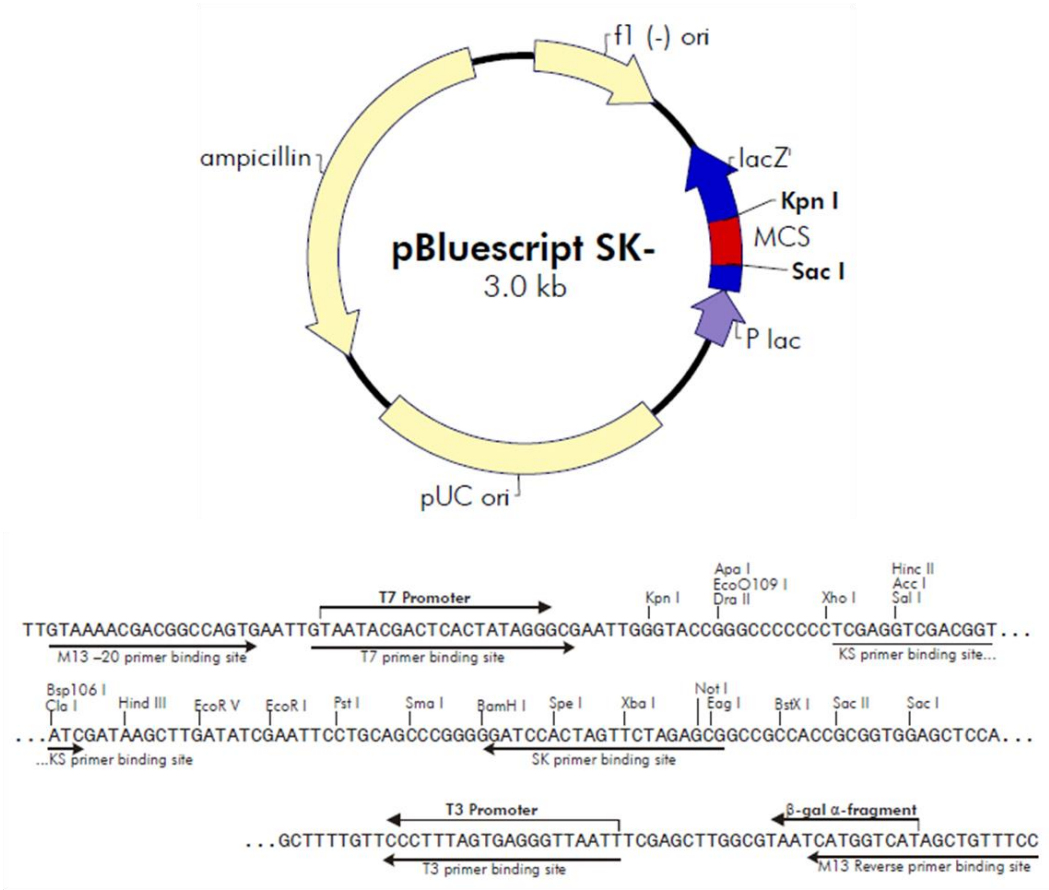
Za izolaciju genomske DNA (gDNA) korišteno je tkivo škrga jedne dagnje *M. galloprovincialis*. Postupak izolacije genomske DNA proveden je koristeći komercijalni komplet Genomic Tip-100G kit (Qiagen), prema uputama proizvođača. Tkivo škrga (100 mg) prebačeno je u 30 ml tubicu sa 9,5 ml pufera za lizu (G2) sastava: 800mM guanidine HCl, 30mM Tris-Cl, pH 8,0, 30 mM EDTA, pH 8.0, 5% Tween-20, 0.5% Triton X-100 i dodatak 19 µL RNase A (Sigma, 100 mg/mL). U puferu za lizu tkivo je homogenizirano na ledu koristeći mehanički štapni homogenizator (Tehnica Železniki). Homogenatu je dodano 500 µL Proteinaze K (Merck, 20 mg/ml) te kratko promiješano na tresilici i inkubirano na 50°C (u vodenoj kupelji) 3h dok lizat nije postao bistar. Kako bi uklonili sitne komadiće nerazgrađenog tkiva koji bi u daljnjem postupku mogli začepiti kolonicu, lizat je centrifugiran na 5000g 10 min na +4°C. Supernatant je odmah prebačen na ranije ekvilibrirane kolonice (QBT pufer; 750 mM NaCl, 50 mM MOPS, pH 7,0, 15% isopropanol, 0,15% Triton X-100) za izolaciju gDNA i odmah započinje gravitacijom potaknuti protok lizata kroz kolonicu. DNA se vezala na matriks (smolu) kolonice (u uvjetima niske koncentracije soli pufera za lizu), a ostali stanični dijelovi (proteini, ugljikohidrati, metaboliti i sl.) su prošli kroz kolonicu. Nakon ispiranja kolonica (QC pufer; 1,0 M NaCl, 50 mM MOPS, pH 7,0, 15% isopropanol), zbog uklanjanja mogućih preostalih nečistoća, DNA je eluirana sa 5 mL pufera visoke koncentracije soli (QF pufer; 1,25 M NaCl, 50 mM Tris-Cl, pH 8,5, 15% isopropanol), a zatim i precipitirana dodatkom 3,5 mL isopropanola (100%) u eluat. Centrifugiranjem na 5000g i +4°C u trajanju od 15 min, oblikovao se „staklasti“ talog DNA. Talog DNA ispran je sa 2 mL hladnog etanola (70%), a nakon centrifugiranja na 5000g i +4°C u trajanju od 10 min, talog DNA resuspendiran je u 1 mL TE pufera (10 mM Tris-Cl, pH 8,0, 1 mM EDTA, pH 8,0). Veličina i kvaliteta izolirane genomske DNA provjerena je elektroforezom u agaroznom gelu.

3.3. Izolacija ukupne RNA i priprema cDNA biblioteke

Za izolaciju ukupne RNA korišteno je tkivo škrga dagnje. Tkivo škrga (15 mg) jedne dagnje zamrznuto je i zdrobljeno u tekućem dušiku te je postupak izolacije ukupne RNA proveden koristeći komercijalni komplet „RNeasy Protect Kit“ (Qiagen), prema uputama proizvođača. Postupak uključuje stabilizaciju ukupne RNA reagensom RNA*later*, denaturaciju puferom sa gvanidin tiocijanatom (čime su inaktivirane i RNaze) te izolaciju ukupne RNA na selektivnim silikatnim kolonicama. RNA molekule duže od 200 nukleotida (kao što su mRNA) uspješnije se vežu na kolonice od manjih RNA molekula (kao što su 5.8S rRNA, 5S rRNA i tRNA). RNA je sa kolonica eluirana u 40 µL vode slobodne od RNaza (engl. *RNase-Free Water*). Tako pripremljena RNA poslana je u servis ustanove Macrogen (Macrogen Inc., Korea) na konstrukciju cDNA biblioteke. Konstrukcija cDNA biblioteke provedena je prema uputama iz komercijalnog kompleta „Uni-ZAP^{TM} XR-Cloning Kit“ (Stratagene). Prvi korak u konstrukciji cDNA biblioteke je prepisivanje izolirane mRNA u cDNA pomoću enzima reverzne transkriptaze. cDNA se zatim direktno klonira u E*co*RI/X*ho*I mjesta lambda Uni-ZAP^{TM} XR vektora (Stratagene) (Slika 6). Raspon veličina inserta (cDNA) bio je 1-3 kb. Rekombinantni vektori pakiraju se u čestice bakteriofaga lambda pomoću „Gigapack III Gold“ ekstrakta koji sadrži sve komponente bakteriofaga lambda. Na ovaj način pripremljena je primarna biblioteka cDNA koju je nužno umnožiti jer su primarni bakteriofagi nestabilni. Umnožena cDNA biblioteka pohrani se na -80°C. Zbog jednostavnije karakterizacije inserta (cDNA) u metodama koje bi koristile cDNA biblioteku kao kalup (umnažanje DNA, određivanje slijeda nukleotida i sl.) iz rekombinantnih lambda Uni-ZAP^{TM} XR vektora provedeno je brzo izrezivanje *in vivo* fagmida pBluescript SK (-) (Slika 7). Izrezanim fagmidom inficirane su stanice *E.coli* soja SOLR, odgovarajućim medijem potaknuto njihovo razmnožavanje te na taj način uspostavljena cDNA biblioteka fagmida.



Slika 6. Skica Uni-ZAP^{TM} XR insercijskog vektora



Slika 7. Skica pBluescript SK- fagmidnog vektora i slijed nukleotida višestrukog mjesta za kloniranje.

3.4. Umnažanje DNA

Sljedovi kodirajuće DNA prisutni u biblioteci kao i fragmenti genomske DNA gena p63 umnoženi su lančanom reakcijom polimeraze (engl. *polymerase chain reaction*, PCR). Korištene su početnice sintetizirane prema kodirajućim DNA sljedovima gena porodice p53 roda *Mytilus* dostupne u GenBank bazi podataka (NCBI) i početnice prema sljedovima p63 gena *Mytilus galloprovincialis* (osobito intronske regije) koji su rezultat izrade ovog rada. Sve korištene početnice navedene su u Tablici 1. U biblioteci kodirajuće DNA smo očekivali zastupljenost više mogućih ekspresijskih izoformi gena pa su korištene razne kombinacije početnica kako bi uspješno umnožili sve izoforme. Također, cjeloviti slijed gena porodice p53 prevelik je da bi ga umnožili u jednoj PCR reakciji pa su korištene razne kombinacije početnica kako bi se dobili preklapajući fragmenti. DNA je umnožena u PCR uređaju C1000^{TM} Thermal Cycler (Bio-Rad). Ukupni volumen reakcijske otopine za PCR je bio 50 µL. Sastav PCR otopine konačnog volumena 50 µL: 1x pufer za PCR (Fermentas), 2 mM $MgCl_2$, 0.2 mM svakog nukleotida, 0.3 µM svake početnice, 1.25 U DreamTaq polimeraze (Fermentas) i 4 µL gDNA / 2 µL cDNA. PCR program se sastojao od početne denaturacije na 95°C 3 min nakon čega je slijedilo 35 ciklusa umnažanja (95°C 30 sec., 58-62°C (ovisno o paru početnica) 30 s, 72°C 1- 4 min (ovisno o očekivanoj dužini umnažanog fragmenta (produkta)) te finalna ekstenzija od 15 min na 72°C.

Tablica 1 Nazivi i pripadajući sljedovi početnica korištenih za umnažanje genomske i kodirajuće DNA gena porodice p53 dagnje *Mytilus galloprovincialis*.

Početnica	5' slijed 3'
MetE1a	AGTGCCAATATGGAACTCCCGT
StartF1	CCATAATGTCACAAGCTTCAGTT
E3R	GACTTCTCCCAGAGTATTCCAT
dNF	GAGAGAACTGGATTTACAACCTA
27R	CTCGGTATTGTTTACTTTCATTTG
N20-1	TACGTCAGAATGGCAACTACTTG
3'2R	TGGTGGGAGAGCTGCCTTCT
StopF	CGACCAGGAATCAAATCAAGAAC
K-boxF	GTGATATGTCTTTGAAATGTGCT
K-boxR	AGCACATTTCAAAGACATATCACAT
Int10R	GATTTTACATTAGGCCTTTCTGCC
Int12F	ATGACAACTTTTCAAAAGGCTTG
E12F	CCTTGGTTTATCAGCCTACATC
E12R	CGATGTAGGCTGATAAACCAA
E13F	GAGTCTATTACCATTAGTGATTC
E13R	GAATCACTAATGGTAATAGACTC
MetUTRev	ATGGCAGTTACTAATGACAACAA
MetUTRev2	TGAATTACAATGGTGGCAATGG
T3	GCAATTAACCCTCACTAAAGGGA
T7	TAATACGACTCACTATAGGG
pUC-M13f	CGCCAGGGTTTTCCCAGTCACGAC
pUC-M13r	TCACACAGGAAACAGCTATGAC

3.5. Elektroforeza DNA

Elektroforeza DNA u agaroznom gelu se koristi za procjenu veličine, količine i kvalitete DNA u uzorku. Ovisno o tipu eksperimenta, DNA fragmenti su razdvojeni na agaroznom gelu različitih koncentracija (0,5% - 1%) u TAE puferu (40mM Tris, 20mM octena kiselina, 1mM EDTA) pri konstantnom naponu od 3-4 V/cm. Gelovi su potom inkubirani u otopini etidij-bromida koncentracije 2 μg/mL u trajanju od 10 minuta te isprani u destiliranoj vodi također 10 minuta. Gelovi su pod UV svjetlom (312 nm) pregledani i fotografirani digitalnim sistemom za dokumentaciju. Veličina fragmenta na gelu je utvrđivana usporedbom sa markerom GeneRuler^{TM} 1 kb DNA Ladder (Fermentas), raspona veličina od 250 do 10 000 pb.

3.6. Izolacija iz gela i pročišćavanje PCR produkata

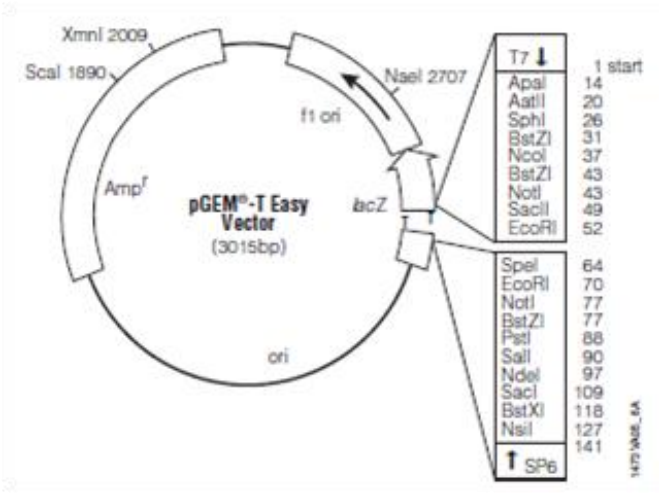
Pregledom gela pod UV svjetlom skalpelom su izrezani komadići gela (DNA vrpce) koji sadrže željene DNA fragmente (PCR produkte). Za pročišćavanje DNA fragmenta korišten je komercijalni komplet MinElute Gel Extraction Kit (Qiagen) prema uputama proizvođača. Metoda se temelji na pročišćavanju DNA ionskom izmjenom na silikatnim kolonama. Fragmenti DNA su sa kolona isprani sa 10 μL pufera (10 mM Tris-Cl, pH 8.5) u kojem su i pohranjeni na +4°C.

3.7. Ligacija u plazmidni vektor i transformacija bakterijskih stanica

Za ligaciju PCR produkata u plazmidni vektor korišten je komercijalni komplet „pGEM-T Easy Vector System“ (Promega). Vektor pGEM-T Easy (Slika 8) je lineariziran i sa po jednim „visećim“ timidinom na oba kraja nukleotidnog slijeda. Timidini sprječavaju recirkularizaciju vektora bez ugradnje inserta, a kompatibilni su i sa adeninima koje je DreamTaq polimeraza dodala na krajeve PCR produkta u koraku finalne ekstenzije 15 min na 72°C te na taj način omogućava uspješnu ugradnju inserta (PCR produkt) u pGEM-T Easy vektor uz pomoć T4 DNA Ligaze. Sastav ligacijske smjese: 2X ligacijski pufer 5 μL, T4 DNA ligaza 1 μL (3 Weiss U), pGEM-T Easy vektor 1 μL (50ng), pročišćeni PCR produkt u količini da je zadovoljen 3:1 molarni omjer insert:vektor te voda slobodna od nukleaza (engl. *Nuclease-Free Water*) do konačnog volumena od 10 μL. Ligacijske smjese su inkubirane 2 sata na sobnoj temperaturi. Insert se ligacijom kovalentno veže u vektor, a vektor time poprima kružni oblik i sposoban je za *in vivo* replikaciju u bakterijskoj kulturi.

Za transformaciju je korišteno 50 µL DH5α kemijski kompetentnih bakterijskih stanica (Invitrogen) u koje je dodano 4 µL ligacijske smjese i inkubirano 30 minuta na ledu. Slijedi temperaturni šok 20 sekundi na 42°C, kratka inkubacija (2 minute) na ledu te dodatak 950 µL SOC hranjivog medija i inkubacija 1 sat u tresilici na 37°C.

Transformirane bakterijske stanice nasade se na LB ploče sa ampicilinom (finalna koncentracija 100 µg/mL) i premazane sa 20 µL X-gal (50 mg/mL) i 100 µL IPTG (24 mg/mL) za selekciju (plavo-bijeli probir) pozitivnih klonova tj. klonova sa ugrađenim insertom. Uzgoj je trajao preko noći na 37°C. Pozitivni klonovi potvrđeni su PCR probiranjem korištenjem vektor specifičnih početnica (pUC/M13f i pUC/M13r, Tablica 1).



Slika 8. Skica pGEM-T Easy plazmidnog vektora.

3.8. Izolacija plazmidne DNA

U svrhu izolacije plazmidne DNA, pozitivni klonovi nasađeni su svaki u po 2 mL tekućeg LB medija sa ampicilinom (finalna kocentracija 100 µg/mL) i uzgojene preko noći na 37°C. Iz takvih uzgojenih, zasićenih bakterijskih kultura izolirana je plazmidna DNA koristeći komercijalni komplet „QIAprep Spin Miniprep Kit“ (Qiagen) prema uputam proizvođača. Metoda se temelji na nježnoj lizi stanica uz prisustvo RNaze A. Time se relativno mala plazmidna DNA oslobađa u otopinu dok bakterijska DNA zaostaje u stanici. Istovremeno se razgrađuju prisutne RNA molekule. Stanični debris (zajedno sa bakterijskom DNA) se istaloži i centrifugira, a plazmidna DNA se iz preostale otopine izolira ionskom izmjenom na silikatnim kolonicama. Konačno, plazmidna DNA eluirana je sa 50 µL EB pufera (10 mM

Tris-Cl, pH 8,5). Koncentracija izolirane plazmidne DNA procijenjena je elektroforezom na 1%-tnom agaroznom gelu.

3.9. Određivanje primarne strukture DNA (sekvenciranje)

Slijed nukleotida kloniranih DNA fragmenata određen je u servisima ustanove Macrogen (Macrogen Inc., Korea) na ABI3730XL DNA Analyzer uređaju. Prilikom određivanja nukleotidnog slijeda (sekvenciranje) korištene su plazmidne početnice (pUC/M13f i pUC/M13r, Tablica 1) ili one specifično konstruirane za pojedine uzorke.

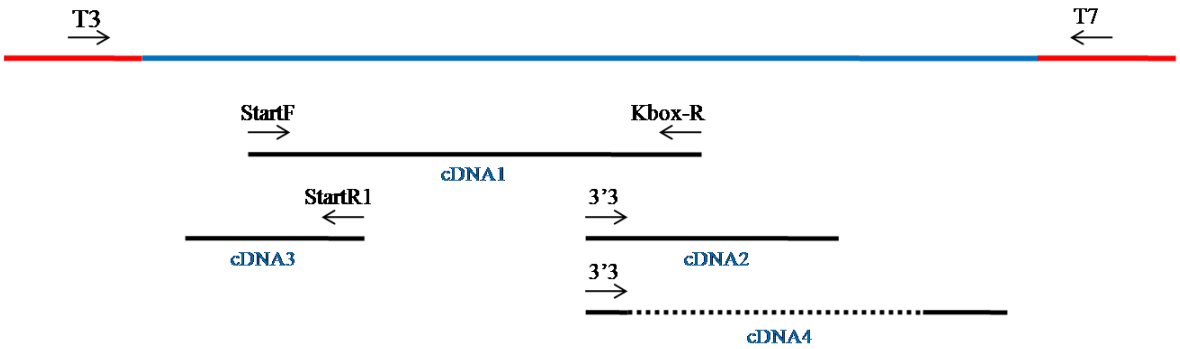
3.10. Računalna analiza sljedova DNA i proteina

Svi sekvencirani sljedovi plazmidne DNA obrađeni su najprije na način da su iz daljnje analize izuzeti sljedovi vektora uzvodno ili nizvodno od slijeda inserta (PCR produkta). Takva početna obrada nukleotidnih sljedova provedena je programskim paketom Lasergene 7 (DNASTAR Inc.). Programi koji su sastavni dio ovih paketa poslužili su za konstrukciju početnica (PrimerSelect), sravnjivanje sljedova i određivanje konsenzus slijeda za pojedine skupine sljedova (SeqMan) te obilježavanje pojedinih značajnih regija (SeqBuilder). Uz Lasergene 7, korišten je i program Geneious 7.1.7. (Biomatters Ltd.) za obradu i prikaz nukleotidnih i proteinskih sljedova. Uz programe ClustalX (Thompson i sur., 1997) i GeneDoc (Nicholas, 1997), program Geneious 7.1.7. (Biomatters Ltd.) je korišten i za višestruka sravnjenja proteinskih sljedova, filogenetske analize i prikaz dobivenih rezultata. Za filogenetsku analizu korišteno je sravnjenje samo najočuvanije središnje regije (DBD, prisutna kod svih izoformi) proteina porodice p53 različitih organizama te metoda sparivanja susjednih grana (eng. *neighbour joining*, NJ). Vjerodostojnosti filogenetskih stabala su provjerene metodom samoučitavanja (engl. *bootstraping*). Konstruirana su jednoglasna stabla (eng. *consensus tree*). Programima dostupnim na mreži provedena su pretraživanja sljedova u odnosu na različite substrukture kao što su ponavljanja (Tandem Repeats Finder, (Benson, 1999)), očuvane domene (NCBI Batch CD-Search Tool, (Marchler-Bauer i sur., 2011)) i motivi (RegRNA 2.0., (Chang i sur., 2013)). Za sravnjenja genomskih sa kodirajućim DNA sljedovima korišten je program Splign (Kapustin i sur., 2008) dostupan na mreži. Svi dobiveni sljedovi uspoređeni su sa onima dostupnim u NCBI GenBank bazi podataka (http://www.ncbi.nlm.nih.gov/) koristeći BLAST servis (Altschul i sur., 1990). Sekvencirani genomi također su pretraživani BLAST servisom.

4. REZULTATI

4.1. Sljedovi kodirajuće DNA porodice p53 *Mytilus galloprovincialis*

Pretraživanje cDNA biblioteke, koristeći metodu lančane reakcije polimerazom (PCR) i specifično dizajnirane početnice, rezultiralo je s četiri nepotpuna cDNA slijeda nazvana cDNA1-4 (Slika 9).



Slika 9. Shema pretraživanja cDNA bibliteke dagnje *Mytilus galloprovincialis*. Umnoženi kodirajući sljedovi gena porodice p53 označeni su plavom bojom (cDNA1-4), a slijed vektora cDNA biblioteke crvenom. Strjelice označavaju smještaj i usmjerenje korištenih početnica.

Slijed cDNA1 je najduži i sadrži 2013 nukleotida (nt); umnožen je koristeći početnice StartF i Kbox-R. Prevođenjem ovog cDNA slijeda dobiva se neprekinuti protein dužine 428 aminokiselina. Otvoreni okvir čitanja određen je startnim kodonom (ATG) centriranim na 33-35 nt i TGA stop kodonom centriranim na 1287-1289 nt cDNA1 slijeda. Slijed cDNA1 sadrži tako dio 5' nekodirajuće regije (5'UTR) od 32 nukleotida, otvoreni okvir čitanja od 1257 nukleotida te 3' nekodirajuću regiju (3'UTR) od 724 nukleotida. 3' UTR cDNA1 ne završava poli-A repom pa je cDNA1 na 3' kraju nepotpuna. Detaljnijom analizom 3' UTR cDNA1 slijeda utvrđena je prisutnost poznatih očuvanih funkcionalnih motiva nekodirajućih regija. Na poziciji 1989-1996 nt smješten je K-box motiv (cTGTGATa), a mjesta vezanja proteina porodice Musashi pronađena su na pozicijama 1571-1576 nt i 1933-1937 nt (Tablica 2, Slika 10). BLAST analiza nukleotidnog slijeda cDNA1 otkriva vrlo visoku sličnost (97 i 94% identiteta na 100% pokrivenosti) sa cjelovitim „p53-like“ kodirajućim sljedovima *Mytilus edulis* i *M. trossulus* (AY579472 i AY611471) (Slika10). BLASTx u proteinskom slijedu cDNA1 potvrđuje prisustvo očuvanih domena porodice p53; transaktivacijsku (TA), DNA vezujuću (DBD) i oligomerizacijsku (OD). Analiza slijeda (418 aminokiselina) otvorenog

okvira čitanja cDNA1 *Batch Web CD-Search Tool* (NCBI) alatom utvrdila je da TA domena obuhvaća aminokiseline 1-16, DBD 126-306, a OD 339-374.

Tablica 2 Popis pronađenih očuvanih funkcionalnih motiva 3'UTR slijeda cDNA1 koristeći RegRNA 2.0. analizu.

Naziv motiva	Pozicija (nt)	Dužina (nt)	Takson	Slijed
K-Box (KB)	1989~1996	8	3UTR Metazoa	atgtgata
Protein porodice Musashi (MBE)	1571~1576	6	3UTR Metazoa	gttagt
Protein porodice Musashi (MBE)	1933~1937	5	3UTR Metazoa	atagt

Slijed cDNA2 umnožen je koristeći početnice 3'3 i T7. Početnica T7 specifična je za vektor cDNA biblioteke, a 3'3 za kodirajuće sljedove porodice p53 roda *Mytilus*. Dužina cDNA2 slijeda je 1904 nukleotida i prvih 916 nukleotida 5' kraja preklapajuće je (100% identičnost) s 3' krajem slijeda cDNA1. Preostalih 178 nukleotida slijeda cDNA2 moguće je produženje 3'UTR regije gena s poli-A repom na kraju slijeda. Poliadenilacijski signal određen je na poziciji 1061-1094 nt. Pronađeno je još jedno mjesto vezanja proteina porodice Musashi (ATAGT) na poziciji 1023-1027 nt slijeda cDNA2 (Tablica 3, Slika10). BLAST analiza nukleotidnog slijeda cDNA2 otkriva vrlo visoku sličnost (94% identiteta) s 3' krajevima cjelovitih „p53-like“ kodirajućih sljedova *Mytilus edulis* i *M. trossulus* (AY579472 i AY611471).

Tablica 3 Popis pronađenih očuvanih funkcionalnih motiva 3'UTR slijeda cDNA2 koristeći RegRNA 2.0. analizu.

Naziv motiva	Pozicija (nt)	Dužina (nt)	Takson	Slijed
Protein porodice Musashi (MBE)	1023~1027	5	3UTR Metazoa	atagt
Poliadenilacijski signal (PAS)	1061~1067	6	3UTR Eukarioti	aataaa

Slijed cDNA3 dug je 308 nukleotida. Umnožen je iz cDNA biblioteke koristeći početnice T3 i StartR1. Početnica T3 specifična je za vektor cDNA biblioteke, a StartR za kodirajuće sljedove porodice p53 roda *Mytilus*. Posljednja 184 nukleotida 3' kraja cDNA3 preklapaju se (100% identičnost) sa 5'krajem cDNA1 dok početnih 124 nukleotida cDNA3 slijeda predstavlja produženje cDNA1 na 5' kraju (Slika10). BLAST pretraga pronalazi opet

očekivano vrlo visoku sličnost cDNA3 slijeda sa 5' krajevima kodirajućih sljedova porodice p53 roda *Mytilus* u GenBank bazi. Ipak, početnih 77 nukleotida cDNA3 ne pronalazi sličnost sa nijednim od tih sljedova u bazi kao ni značajniju sličnost sa nekim drugim nukleotidnim (BLASTn) ili proteinskim (BLASTx) slijedom u GenBank bazi. Detaljnijom analizom cDNA3 slijeda pronađeno je 6 očuvanih motiva (mjesta vezanja) regulatornih proteina čovjeka (Tablica 4, Slika 10) smještenih većinom u početnih stotinjak nukleotida 5'kraja cDNA3 slijeda. Regulatorni proteini pronađenih motiva funkcionalno su povezani s porodicom p53.

Tablica 4 Popis pronađenih očuvanih motiva regulatornih proteina slijeda cDNA3 koristeći RegRNA 2.0. analizu.

Naziv motiva	Pozicija (nt)	Dužina (nt)	Takson	Slijed
c-Ets	81~94	14	*Homo sapiens*	tggaaagttcactc
Pax-4	14~25	12	*Homo sapiens*	gaagtgtttgtt
IRF-8	43~49	7	*Homo sapiens*	gagtgtt
ING4	71~78	6	*Homo sapiens*	cataat
GATA-5	207~212	6	*Homo sapiens*	acaaga
MEF-2C	98~104	7	*Homo sapiens*	atcaccc

Slijed cDNA4 umnožen je početnicama 3'3 i T7 kao i slijed cDNA2. Ukupna dužina slijeda cDNA4 je 447 nukleotida, a početnih 48 nukleotida 5'kraja preklapa se sa sljedovima cDNA1 (1145-1192 nt) i cDNA2 (49-96 nt). U preostalih 399 nukleotida cDNA4 slijed je potpuno različit od cDNA1 i 2. BLASTn pronalazi sličnost tog 3' kraja cDNA4 (99% i 96% identiteta) sa 3' krajevima „p63/73-like“ sljedova roda *Mytilus* (DQ865150-DQ865153, DQ060435-DQ060438) u GenBank bazi (Slika 10). BLASTx otkriva da cDNA4 kodira za SAM domenu porodice p53.

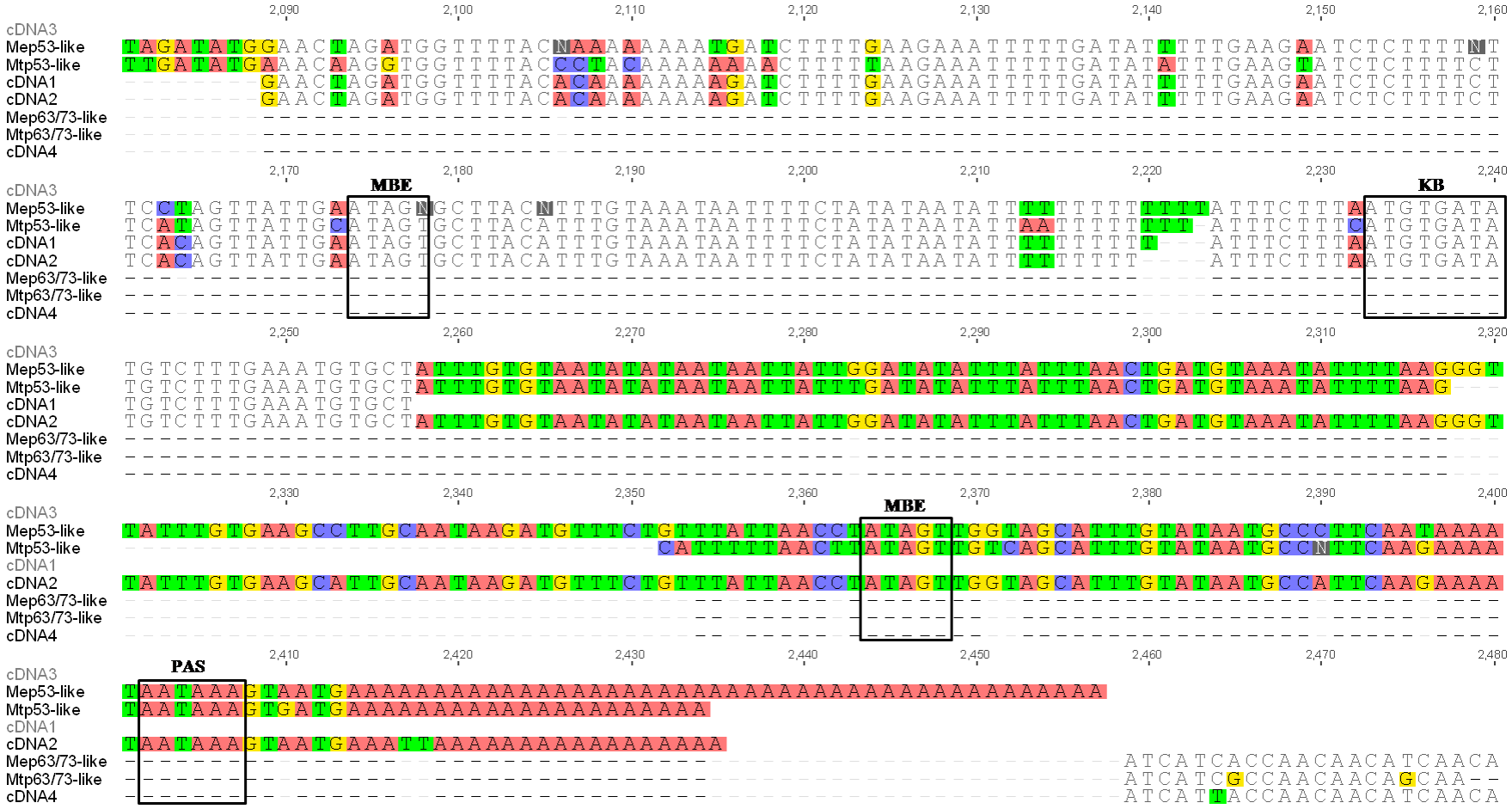
```
                1         10        20        30        40        50        60        70        80
cDNA3                                                                                   GCTTCCTTCAT
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like ATGCCGTCTCAAATGAATGGGGATTTACACAATACACCTCAAGTGCCAATATGGAACTCCCGTCCAGATCCAAGACTGAA
Mtp63/73-like             ATGAATGGGGATTTACACAATACACCTCAAGTGCCAATATGGAACTCCCGTCCAGATCCAAGACTGAA
cDNA4
                    Pax-4 90        100       110  IRF-8 120       130       140 ING4  150   Ets 160
cDNA3         TCGAAGTGTTTGTTTATATCTAGTATGGTTTGAGTGTTTTATCATTATCTTTTCATGTTCATAATCATTTGGAAAGTTCA
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like TTATCTGCGTTCGCAAGGATATAATAATTACAACATGCCGTCACATAATTATAATCCAGGATGCAGATTTGGAAAGTTCA
Mtp63/73-like TTATCTGCGTTCTCAAGGATACAATAATAACAGCATGCCGTCACATAATTATAATCCAGGATGCAGATTTGGAAAATTCA
cDNA4
                        170 MEF-2C 180       190       200       210       220       230       240
cDNA3         CTCATCATCACCCCATAATGTCACAAGCTTCAGTTTCAACTACATGCACACCATCAGGCCCACCAATGAGCCAGGAAACA
Mep53-like                CCATAATGTCACAAGCTTCAGTTTCAACTACATGCACACCATCAGGCCCACCAATGAGCCAGGAAACA
Mtp53-like                CCATAATGTCACAAGCTTCAGTTTCAACTACATGCACACCATCAGGCCCACCAATGAGCCAAGAAACA
cDNA1                                          TTTCAACTACATGCACACCATCAGGCCCACCAATGAGCCAGGAAACA
cDNA2
Mep63/73-like CTCATCATCACCCCATAATGTCACAAGCTTCAGTTTCAACTACATGCACACCATCAGGCCCACCAATGAGCCAGGAAACA
Mtp63/73-like CTCATCATCACCCCATAATGTCACAAGCTTCAGTTTCAACTACATGCACACCATCAGGCCCACCAATGAGCCAAGAAACA
cDNA4
                        250       260       270       280       290       300       310       320
cDNA3         TTTGAATATTTATGGAATACTCTGGGAGAAGTCACACAAGAAGGGGGATACACAAACATCACATCTAAAGAATCAATCGA
Mep53-like    TTTGAATATTTATGGAATACTCTGGGAGAGGTCACACAAGAAGGGGGATACACAAACATCACATCTAAAGAATCAATCGA
Mtp53-like    TTTGAATACTTATGGAATACTCTGGGAGAAGTCACACAAGAAGGGGGATACACAAACATCACATCTAAAGAATCAATCGA
cDNA1         TTTGAATATTTATGGAATACTCTGGGAGAAGTCACACAAGAAGGGGGATACACAAACATCACATCTAAAGAATCAATCGA
cDNA2
Mep63/73-like TTTGAATATTTATGGAATACTCTGGGAGAGGTCACACAAGAAGGGGGATACACAAACATCACATCTAAAGAATCAATCGA
Mtp63/73-like TTTGAATACTTATGGAATACTCTGGGAGAAGTCACACAAGAAGGGGGATACACAAACATCACATCTAAAGAATCAATCGA
cDNA4
                        330       340       350       360       370       380       390       400
cDNA3         CTATGCATTCAGTGAAGCTGAGGATGAAACCAGTATATCTGTAGAGAAATACAGGAT
Mep53-like    CTATGCATTCAGTGAAGCTGAGGATGAAACCAGTATATCTGTTGAGAAATACAGGATTACATCCAATGATTCCATTTCAG
Mtp53-like    CTATGCATTTAGTGAAGCTGAGGATGAAACCAGCATATCTGTTGAGAAATACAGAATTACATCCAATGACTCCATTTCAG
cDNA1         CTATGCATTCAGTGAAGCTGAGGATGAAACCAGTATATCTGTAGAGAAATACAGGATTACATCCAATGATTCCATTTCAG
cDNA2
Mep63/73-like CTATGCATTCAGTGAAGCTGAGGATGAAACCAGTATATCTGTTGAGAAATACAGGATTACATCCAATGATTCCATTTCAG
Mtp63/73-like CTATGCATTTAGTGAAGCTGAGGATGAAACCAGTATATCTGTTGAGAAATACAGAATTACATCCAATGACTCCATTTCAG
cDNA4
                        410       420       430       440       450       460       470       480
cDNA3
Mep53-like    ACTTGTTGAATCCTATCATAGGTCAAACGACTACAGCTAGCTCTATGAGCCCTGACAGTCAGACCAATATCATCGGTAGC
Mtp53-like    ACTTGTTGAATCCTATCATAGGTCAAACGACCTCAGCGAGCTCTATGAGCCCTGACAGTCAGACCAACATCATCGGTAGC
cDNA1         ACTTGTTGAATCCTATCATAGGTCAAACGACTACAGCTAGCTCTATGAGCCCTGACAGTCAGACCAATATCATCGGTAGC
cDNA2
Mep63/73-like ACTTGTTGAATCCTATCATAGGTCAAACGACTACAGCTAGCTCTATGAGCCCTGACAGTCAGACCAATATCATCGGTAGC
Mtp63/73-like ACTTGTTGAATCCTATCATAGGTCAAACGACCTCAGCGAGCTCTATGAGCCCTGACAGTCAGACCAACATCATCGGTAGC
cDNA4
```

```
                        490       500       510       520       530       540       550       560
cDNA3
Mep53-like     TCTGCATCAAGTCCTTACAATGATACAATCACATCACCCCCACCTTATTCGCCTCATACTAGTATGCAGTCACCTATCCC
Mtp53-like     TCTGCATCAAGTCCTTACAATGATACAATCACATCACCTCCACCTTATTCGCCTCATACTAGTATGCAGTCGCCTATACC
cDNA1          TCTGCATCAAGTCCTTACAATGATACAATCACATCACCCCCACCTTATTCGCCTCATACTAGTATGCAGTCACCTATACC
cDNA2
Mep63/73-like  TCTGCATCAAGTCCTTACAATGATACAATCACATCACCCCCACCTTACTCGCCTCATACTAGTATGCAGTCACCTATACC
Mtp63/73-like  TCTGCATCAAGTCCTTACAATGATACAATCACATCACCTCCACCTTATTCGCCTCATACTAGTATGCAGTCGCCTATACC
cDNA4
                        570       580       590       600       610       620       630       640
cDNA3
Mep53-like     CTCTGTGCCATCAAACACAGACTATCCCGGCGATTATGGGTTCACCATTTCATTTTCTCAGCCATCAAAAGAGACCAAGT
Mtp53-like     CTCTGTGCCATCAAACACTGACTATCCAGGAGATTATGGCTTCACCATTTCATTTTCTCAGCCATCAAAAGAGACAAAGT
cDNA1          CTCTGTGCCATCAAACACAGACTATCCCGGCGATTATGGGTTCACCATTTCATTTTCTCAGCCATCAAAAGAGACAAAGT
cDNA2
Mep63/73-like  CTCTGTGCCATCGAACACAGACTATCCCGGCGATTATGGGTTCACCATTTCATTTTCTCAGCCATCAAAAGAGACAAAGT
Mtp63/73-like  CTCTGTGCCATCAAACACTGACTATCCAGGAGATTATGGCTTCACCATTTCATTTTCTCAGCCATCAAAAGAGACAAAGT
cDNA4
                        650       660       670       680       690       700       710       720
cDNA3
Mep53-like     CTACCACATGGACATACTCAGAGTCACTTAAGAAATTATACGTCAGAATGGCAACAACTTGCCCAATCCGATTTAAATGT
Mtp53-like     CTACCACATGGACATACTCAGAGTCACTTAAGAAATTATACGTCAGAATGGCAACAACTTGCCCAATCCGATTTAAATGT
cDNA1          CTACCACATGGACATACTCAGAGTCACTTAAGAAATTATACGTCAGAATGGCAACAACTTGCCCAATCCGGTTTAAGTGT
cDNA2
Mep63/73-like  CTACCACATGGACATACTCAGAGTCACTTAAGAAATTATACGTCAGAATGGCAACAACTTGCCCAATCCGATTTAAATGT
Mtp63/73-like  CTACCACATGGACATACTCAGAGTCACTTAAGAAATTATACGTCAGAATGGCAACAACTTGCCCAATCCGATTTAAATGT
cDNA4
                        730       740       750       760       770       780       790       800
cDNA3
Mep53-like     TTAAGACAGCCCCCACAGGGATGTGTTATTCGTGCAATGCCAATATTCATGAAACCTGAACATGTCCAAGAACCTGTAAA
Mtp53-like     TTGAGACAGCCCCCACAGGGATGTGTAATTCGTGCAATGCCAATATTCATGAAACCTGAACATGTCCAAGAACCTGTAAA
cDNA1          TTGAGACAGCCTCCACAAGGATGTGTTATTCGTGCAATGCCAATATTCATGAAACCCGAACATGTCCAAGAACCTGTAAA
cDNA2
Mep63/73-like  TTAAGACAGCCCCCACAGGGATGTGTTATTCGTGCAATGCCAATATTCATGAAACCTGAACATGTCCAAGAACCTGTAAA
Mtp63/73-like  TTGAGACAGCCCCCACAGGGATGTGTAATTCGTGCAATGCCAATATTCATGAAACCTGAACATGTCCAAGAACCTGTAAA
cDNA4
                        810       820       830       840       850       860       870       880
cDNA3
Mep53-like     AAGATGCCCAAATCATGCCACATCTAAAGAGCACAATGAAAATCATCCAGCTCCAACACATTTATGTCGATGTGAGCACA
Mtp53-like     AAGATGCCCAAATCATGCCACATCTAAAGAGCACAATGAAAATCATCCAGCCCCAACACATTTATGTCGATGTGAGCACA
cDNA1          AAGATGCCCAAATCATGCCACATCTAAAGAGCACAATGAAAATCATCCAGCTCCAACACATTTATGTCGATGTGAGCACA
cDNA2
Mep63/73-like  AAGATGCCCAAATCATGCCACATCTAAAGAGCACAATGAAAATCATCCAGCTCCAACACATTTATGTCGATGTGAGCACA
Mtp63/73-like  AAGATGCCCAAATCATGCCACATCTAAAGAGCACAATGAAAATCATCCAGCCCCAACACATTTATGTCGATGTGAGCACA
cDNA4
```

```
                         890       900       910       920       930       940       950       960
cDNA3
Mep53-like     AACTTGCTAAATTTGTTGAAGATCCATATACCAGCCGCCAGAGTGTTCTAATTCCACATGAGATACCTCAAGCTGGCTCA
Mtp53-like     AACTTGCTAAATTTGTTGAAGACCCATACACAAGCCGCCAGAGTGTTCTAATTCCACATGAGATACCTCAAGCTGGCTCA
cDNA1          AACTTGCTAAATTTGTTGAAGATCCATATACCAGCCGCCAGAGTGTTCTAATTCCACACGAGATACCTCAAGCTGGCTCA
cDNA2
Mep63/73-like  AACTTGCTAAATTTGTTGAAGATCCATATACCAGCCGCCAGAGTGTTCTAATTCCACATGAGATACCTCAAGCTGGCTCA
Mtp63/73-like  AACTTGCTAAATTTGTTGAAGACCCATACACAAGCCGCCAGAGTGTTCTAATTCCACATGAGATACCTCAAGCTGGCTCA
cDNA4

                         970       980       990      1,000     1,010     1,020     1,030     1,040
cDNA3
Mep53-like     GAATGGGTCACCAATTTGTTCCAGTTCATGTGCCTGGGGTCATGTGTAGGAGGACCAAACAGAAGGCCTATTCAGATTGT
Mtp53-like     GAATGGGTCACCAATTTGTTCCAGTTCATGTGTCTGGGGTCATGTGTAGGAGGACCAAACAGAAGACCTATTCAGATTGT
cDNA1          GAATGGGTCACCAATTTGTTCCAGTTCATGTGCCTGGGGTCATGTGTAGGAGGACCAAACAGAAGGCCTATTCAAATTGT
cDNA2
Mep63/73-like  GAATGGGTCACCAATTTGTTCCAGTTCATGTGCCTGGGTTCATGTGTAGGAGGACCAAACAGAAGGCCTATTCAGATTGT
Mtp63/73-like  GAATGGGTCACCAATTTGTTCCAGTTCATGTGCCTGGGGTCATGTGTAGGAGGACCAAACAGAAGACCTATTCAGATTGT
cDNA4

                        1,050     1,060     1,070     1,080     1,090     1,100     1,110     1,120
cDNA3
Mep53-like     TCTGACTTTAGAAAAAGATAATCAAGTGCTAGGTAGACGGGCAGTCGAAGTAAGAATCTGTGCCTGTCCTGGGAGAGACA
Mtp53-like     TCTGACTTTAGAAAAAGATAATCAAGTGCTAGGTAGACGTGCCGTAGAAGTGAGAATTTGTGCCTGTCCTGGAAGAGATC
cDNA1          TCTGACTTTAGAAAAAGATAATCAAGTGCTAGGTAGACGGGCAGTCGAAGTAAGAATTTGTGCCTGTCCTGGGAGAGACC
cDNA2
Mep63/73-like  TCTGACTTTAGAAAAAGATAATCAAGTGCTAGGTAGACGGGCAGTGGAAGTTAGAATTTGCGCCTGTCCTGGGAGAGACA
Mtp63/73-like  TCTGACTTTAGAAAAAGATAATCAAGTGCTAGGTAGACGTGCCGTAGAAGTGAGAATTTGTGCCTGTCCTGGAAGAGATC
cDNA4

                        1,130     1,140     1,150     1,160     1,170     1,180     1,190     1,200
cDNA3
Mep53-like     GAAAGGCTGATGAAAAGGCAGCTCTCCCACCATGCAAACAGTCCCCAAAGAAAGGCCAGAAAGTTAATATTATCAATGAA
Mtp53-like     GTAAGGCTGATGAGAAGGCAGCTCTCCCACCATGTAAACAGTCCCCAAAGAAAGGCCAGAAAGTTAATATTATCAATGAA
cDNA1          GAAAGGCTGATGAGAAGGCAGCTCTCCCACCACGCAAACAGTCCCCAAAGAAAGGCCAGAAAGTTAATATTATCAATGAA
cDNA2
Mep63/73-like  GAAAGGCTGATGAAAAGGCAGCTCTCCCACCATGCAAACAGTCCCCAAAGAAAGGCCAGAAAGTTAATATTATCAATGAA
Mtp63/73-like  GTAAGGCTGATGAGAAGGCAGCTCTCCCACCATGTAAACAGTCCCCAAAGAAAGGCCAGAAAGTTAATATTATCAATGAA
cDNA4

                        1,210     1,220     1,230     1,240     1,250     1,260     1,270     1,280
cDNA3
Mep53-like     ATCACTACAGTAACACCAGGAGGCAAAAAGAGGAAAGCAGAAGACGAACCATTCACATTATCTGTACGAGGACGAGAAAA
Mtp53-like     ATCACTACAGTAACACCAGGAGGCAAAAAGAGGAAAGCAGAAGATGAACCCTTCACATTGTCTGTACGAGGACGAGAAAA
cDNA1          ATCACTACAGTAACACCAGGAGGCAAAAAGAGGAAAGCAGAAGACGAACCATTCACATTATCTGTACGAGGACGAGAAAA
cDNA2
Mep63/73-like  ATCACTACAGTAACACCAGGAGGCAAAAAGAGGAAAGCAGAAGACGAACCATTCACATTATCTGTACGAGGACGAGAAAA
Mtp63/73-like  ATCACTACAGTAACACCAGGAGGCAAAAAGAGGAAAGCAGAAGATGAACCCTTCACATTGTCTGTACGAGGACGAGAAAA
cDNA4
```

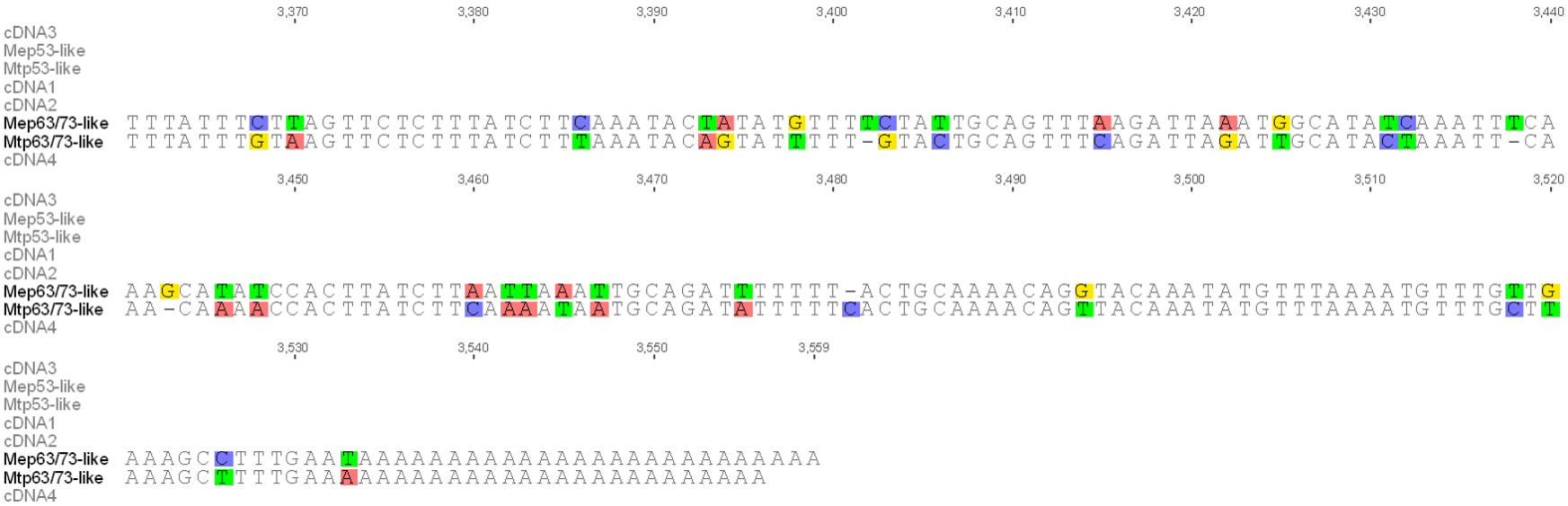
```
                    1,290     1,300     1,310     1,320     1,330     1,340     1,350     1,360
cDNA3
Mep53-like     CTACGAAATTCTGTGTAGACTGAGGGATTCATTGGAACTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAAC
Mtp53-like     CTACGAAATTCTGTGTAGACTGAGGGATTCATTGGAATTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAAC
cDNA1          CTACGAAATTCTGTGTAGACTGAGGGATTCATTGGAACTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAAC
cDNA2                   TCTGTGTAGACTGAGGGATTCATTGGAACTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAAC
Mep63/73-like  CTATGAAATTCTGTGTAGACTGAGGGATTCATTGGAACTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAAC
Mtp63/73-like  CTACGAAATTCTGTGTAGACTGAGGGATTCATTGGAATTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAAC
cDNA4                                                                GAATCAAATTGATGTATACAAAC
                    1,370     1,380     1,390     1,400     1,410     1,420     1,430     1,440
cDNA3
Mep53-like     AGAAACAACTTGATACAAACAGACAGTGGCTGTCGATGATACTGGCCAGGGAGAACAAAAACAAACTGATGAAGAAGGTG
Mtp53-like     AGAAACAACTGGGATACAAATAGACAGTGGCTGTCGATGATACTGGCCAGGGAGAACAAAAACAAACTGATGAAGAAGGTG
cDNA1          AGAAACAACTTGATACAAACAGACAGTGGCTGTCGATGATACTGGCCAGGGAGAACAAAAACAAACTGATGAAGAAGGTG
cDNA2          AGAAACAACTTGATACAAACAGACAGTGGCTGTCGATGATACTGGCCAGGGAGAACAAAAACAAACTGATGAAGAAGGTG
Mep63/73-like  AGAAACAACTTGATACAAACAGACA-----------------------------------------------------
Mtp63/73-like  AGAAACAACTGGGATACAAATAGACA-----------------------------------------------------
cDNA4          AGAAACAACTTGATACAATCAGACA-----------------------------------------------------
                    1,450     1,460     1,470     1,480     1,490     1,500     1,510     1,520
cDNA3
Mep53-like     AAACGACCTCAACATCGACCAGGAATCAAATCAAGAACTTGAAGAGAGAAGTTTCCAATTGGTTCAGGAACATTGAGGAT
Mtp53-like     AAACGACCTCAGCATCGACCAGGAATCAAATCAAGAACTTGAAGAGAGAAGTTTCCAATTGGTTCAGGAACATTGAGGAT
cDNA1          AAACGACCTCAACATCGACCAGGAATCAAATCAAGAACTTGAAGAGAGAAGTTTCCAACTGGTCCAGGAACATTGAGGAT
cDNA2          AAACGACCTCAACATCGACCAGGAATCAAATCAAGAACTTGAAGAGAGAAGTTTCCAACTGGTCCAGGAACATTGAGGAT
Mep63/73-like  ----------------------------------------------------------------------------
Mtp63/73-like  ----------------------------------------------------------------------------
cDNA4          ----------------------------------------------------------------------------
                    1,530     1,540     1,550     1,560     1,570     1,580     1,590     1,600
cDNA3
Mep53-like     ATATAAAGACCCTGCTTTAATGTGGCGCATTCATTTTGGACAATTATAGACATTTATCAAAATCTTTTGGTATACAGTGA
Mtp53-like     ATATAAGGACCTTGCTTTAATGTGGTGCATTCATTT-GGACTATAATAGATATTTATCAAAATCTTTTGGTATACAGTAA
cDNA1          ATATAAGGACCTTGCTTTAATGTGGTGCATTCATTTTGGACAATTATAGACATTTATCAAAATATTTTGGTATACAGTGA
cDNA2          ATATAAGGACCTTGCTTTAATGTGGTGCATTCATTTTGGACAATTATAGACATTTATCAAAATATTTTGGTATACAGTGA
Mep63/73-like  ----------------------------------------------------------------------------
Mtp63/73-like  ----------------------------------------------------------------------------
cDNA4          ----------------------------------------------------------------------------
                    1,610     1,620     1,630     1,640     1,650     1,660     1,670     1,680
cDNA3
Mep53-like     GATCCAGATGACTCCCTGCTCCTGAAAGCAAGGATAATTGATATTTTATAGGACTTATCCCGTCCCCTCCTGCCATTTCC
Mtp53-like     GATCCAGATGCGTCC---------TGAAAGCAAGGATAATTGATATTTTAACGAACTTATCACGTCACCTCCTGCCATTTCA
cDNA1          GATCCAGATGACTCCATGCTCCTGAAAGCAAGGATAATTGATATTTTATAGAACTTATCACGTCACCTCCTGCCATTTCA
cDNA2          GATCCAGATGACTCCATGCTCCTGAAAGCAAGGATAATTGATATTTTATAGAACTTATCACGTCACCTCCTGCCATTTCA
Mep63/73-like  ----------------------------------------------------------------------------
Mtp63/73-like  ----------------------------------------------------------------------------
cDNA4          ----------------------------------------------------------------------------
```

1,690 1,700 1,710 1,720 1,730 1,740 1,750 1,760
cDNA3
Mep53-like CTGATGACCAAGGCAGCAACATATGTGGTCAATTCCTTATCGTTCCATTGGTTGGTGGTGGTGAGGCCCCATATTGGTTT
Mtp53-like CTGATGACCAAGGCAGCAACTAATGTGTTCAAATCATTATCGTTTCATTGGTTGTTGGT----AGGTCTCATATTGGTTT
cDNA1 CTGATGACCAAGGCAGCAACATATGTGTTCAATTCATTATCGTTTCATTGTTTGTTGTTGTTT-AGGTCTCATATTGTTTT
cDNA2 CTGATGACCAAGGCAGCAACATATGTGTTCAATTCATTATCGTTTCATTGTTTGTTGTTGTTT-AGGTCTCATATTGTTTT
Mep63/73-like
Mtp63/73-like
cDNA4
1,770 1,780 1,790 1,800 1,810 1,820 1,830 1,840
MBE
cDNA3
Mep53-like GGTTGTTAGGTTATAAGAGGTTTGAGTCCATAGATATTTATTTTATAGGGCATGGTTTTAGAAATCAAGAATTTT-AGAG
Mtp53-like GGTTGTTAG-TTATAAGAGTTTTGAGGACATAGATATTTATTTTATAGGGCATGGTTTTAGAAATCAAGAATTTTTAGAG
cDNA1 GTTTGTTAG-TTATAAGAGTTTTGAGTACATAGATATTTATTTTATAGGGCATGGTTTTAGAAATTAAGAATTTT-AGAG
cDNA2 GTTTGTTAG-TTATAAGAGTTTTGAGTACATAGATATTTATTTTATAGGGCATGGTTTTAGAAATTAAGAATTTT-AGAG
Mep63/73-like
Mtp63/73-like
cDNA4
1,850 1,860 1,870 1,880 1,890 1,900 1,910 1,920
cDNA3
Mep53-like GATAATTTTATAT--TTTTTTCTTGTTTTTGAATGACATGGATTGTCTTTTAGTGAAAAGCCTTTGAATCTGTTTCTGTA
Mtp53-like GATAAATATATATATTTGTTTCTTGTTTTTGAATGACATTGATTGTCTTTTAGTGAAAAGCATTTGA-TCTGCTTCTATA
cDNA1 GATAATTTTATAT--TTGTTTCTTGTTTTTGAATGACATGGATTGTCTTTTAGTGAAAAGCATTTGAATCTGTTTCAATA
cDNA2 GATAATTTTATAT--TTGTTTCTTGTTTTTGAATGACATGGATTGTCTTTTAGTGAAAAGCATTTGAATCTGTTTCAATA
Mep63/73-like
Mtp63/73-like
cDNA4
1,930 1,940 1,950 1,960 1,970 1,980 1,990 2,000
cDNA3
Mep53-like TGAAAGAATTTTAAGACGAGAGTTCTGCATCTCATGTTGTTTTTGTTTAGATTTTTTTAT------TTAAATTGAATCAAA
Mtp53-like TGAAAGAAATTTAGGGCAAGAGTTNTGAATCTCAAGT--TTTTTTTTGTTTTTTTTTAT------TTAAATATAAGCAAA
cDNA1 TGAAAGAAATTTAAGACAAGAGTTCTGAATCTCATGTTGTTTTTGTTTAGATTTTTTTATTTTTATTTAAACTGAAGCAAA
cDNA2 TGAAAGAAATTTAAGACAAGAGTTCTGAATCTCATGTTGTTTTTGTTTAGATTTTTTTATTTTTATTTAAACTGAAGCAAA
Mep63/73-like
Mtp63/73-like
cDNA4
2,010 2,020 2,030 2,040 2,050 2,060 2,070 2,080
cDNA3
Mep53-like GTCTATCGGCTTTTTATCTTTTGCCTAACAATTTGATATTTATTGAAAATTTATGGTGTTATAGACTTCCCTCCCAAAAT
Mtp53-like GTCTATCAGCTTTTTATCTTTTGCCTAACAATTTGATATTTATTGAAAATTTATGGTGTTATAGACTTACATCNCAAAAT
cDNA1 GTCTATCGGCTTTTTATCTTTTGCCTAACAATTTGATATTTATTG
cDNA2 GTCTATCGGCTTTTTATCTTTTGCCTAACAATTTGATATTTATTG
Mep63/73-like
Mtp63/73-like
cDNA4


cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like
Mtp63/73-like
cDNA4
MBE
KB
PAS

```
                    2,490     2,500     2,510     2,520     2,530     2,540     2,550     2,560
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  GCAAGGGTTGTTACTCTTCCTACTCACGACAACACACCAATAACTATACAGGGTGAAGGTCGGCAGACGACACTACCATT
Mtp63/73-like  ----GGGTTGTTACTCTTCCTACTCACGACAACACACCAATAACTATACAGGGTGAAGGTCGGCAGACGACACTACCGTT
cDNA4          GCAAGGGTTGTTACTCTTCCTACTCACGACAACACACCAATAACTATACACGGTGAAGGTCGGCAGACGACACTACCATT
                    2,570     2,580     2,590     2,600     2,610     2,620     2,630     2,640
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  CACTGCAGACTTAAATGGTCAGGTGACAAGTTCACAAAATGGTGTGGTGGAAAATCATGGAAATATTAAAGAAGAATTGA
Mtp63/73-like  CACTGCAGACTTAAATGGTCAGGTGACAAGTTCACAAAATGGTGTGGTGGAAAATCATGGAAATATTAAAGAAGAACTGA
cDNA4          CACTGCAGACTTAAATGGTCAGGTGACAAGTTCACAAAATGGTGTGGTGGAAAATCATGGAAATATTAAAGAAGAACTGA
                    2,650     2,660     2,670     2,680     2,690     2,700     2,710     2,720
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  TGGCTAATGGGGACCATTCTATTTCAAATTGGTTGACTACCCTTGGTTTATCAGCCTACATCGATAACTTTCATCAACAG
Mtp63/73-like  TGGCCAATGGGGACCATTCAATTTCAAATTGGTTGACTACTCTTGGTTTATCAGCATACATAGATAACTTTCATCAACAG
cDNA4          TGGCTAATGGGGACCATTCTATTTCAAATTGGTTGACTACCCTTGGTTTATCAGCCTACATCGATAACTTTCATCAACAG
                    2,730     2,740     2,750     2,760     2,770     2,780     2,790     2,800
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  AATCTATTTACAATGGAGCAACTCGATGATTTTACTGTTGAGGACTTACAAAAGATGAGAATAGGAACAAGCCACAGAAA
Mtp63/73-like  AATCTGTTTACAATGGAGCAGCTTGATGATTTTACTGTTGAGGACTTGCAAAAGATGAGAATAGGAACAAGCCACAGAAA
cDNA4          AATCTATTTACAATGGAGCAACTCGATGATTTTACTGTTGAGGACTTACAAAAGATGAGAATAGGAACAAGCCACAGAAA
                    2,810     2,820     2,830     2,840     2,850     2,860     2,870     2,880
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  TAAAATCTGGAAAGCTCTAGTAGAATTCCATAGCGAGTCTATTACCATTAGTGATTCACAAAGTCTGCAGAGGGACGTCA
Mtp63/73-like  CAAAATCTGGAAAGCTCTAGTAGAATTCCATAGCGAGTCTATCACCATTAGTGATTCACAAAGTCTGCAGAGGGACGTCA
cDNA4          TAAAATCTGGAAAGCTTTAGTAGAATTCCATAGCGAGTCTATTACCATTAGTGATTC
```

```
                    2,890     2,900     2,910     2,920     2,930     2,940     2,950     2,960
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  GCACTGCATCAACCATTAGTATGACATCACAGGCCAGTATCTCCCAGAATTCTCAAAATTCAACGTATTGTCCGGGATAC
Mtp63/73-like  GCACTGCATCCACCATTAGTATGACATCACAGGCCAGTATCTCCCAGAATTCTCAAAATTCAACGTATTGTCCAGGATAC
cDNA4
                    2,970     2,980     2,990     3,000     3,010     3,020     3,030     3,040
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  TATGAAGTTACCAGATATACATTCAAACATACAGTTTCTTATACGAAAGTGGATGAAAGAAGTCCAAAACGGGCTAAGGT
Mtp63/73-like  TATGAAGTTACCAGATATACATTCAAACATACAGTTTCTTATACGAAAGTGGATGAAAGAAGTCCAAAACGGGCTAAAGT
cDNA4
                    3,050     3,060     3,070     3,080     3,090     3,100     3,110     3,120
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  AGACTGATATGC----TGAAGACTTATGCCACAACCAAACTAGGCTTGATATGATTGGTCAGTCCTTTACTGTGTATTGG
Mtp63/73-like  AGACTGATATGCAATATGAAGACTTATGCCACCATCAAACTAGGCTTAAAATGATTGGTCAGTCCTTTACTGTGTATTGG
cDNA4
                    3,130     3,140     3,150     3,160     3,170     3,180     3,190     3,200
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  ACTGACATATATAGAGCAAATGCAATATGGTGCACAAATCAATGTCTAAATGAAGACATTGTAATCTGATATTTTGTTAG
Mtp63/73-like  ACTGACATATA--GAGCAAATGCAATATGGTGCACAAATCAATGTCTAAATGTAGACATTTTAATCTGATATTTTGTTAG
cDNA4
                    3,210     3,220     3,230     3,240     3,250     3,260     3,270     3,280
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  ACCGCCTGTTTTTAATGCAAATGTTAGATTCTGTCATGTGACCTGTATGGACC--------------------------
Mtp63/73-like  ATTGCCTGTTTTTAATGCAAATATTAGATTCTGTCATGTGACCTGTATGGACCAATTTTTTTAGACAAAATCTTATTTTT
cDNA4
                    3,290     3,300     3,310     3,320     3,330     3,340     3,350     3,360
cDNA3
Mep53-like
Mtp53-like
cDNA1
cDNA2
Mep63/73-like  -----------------------TATTATTGATTTGAAATGGCTATTTATTAAACCGATGATGATATGTAAAAATGTTGG
Mtp63/73-like  TTGTCATGGGACCCCTGTGGACCAATAATTGATTTGAAATGACTTTTATAAAAATGATGATGATGGGTAAAAATGTTGG
cDNA4
```



Slika 10. Višestruko sravnjenje nukleotidnih sljedova cDNA *M. galloprovincialis* (cDNA1-4), cDNA „p53-like“ *M. edulis* (Mep53-like, AY579472) i *M. trossulus* (Mtp53-like, AY611471) te cDNA „p63/73-like“ *M. edulis* (Mep63/73-like, DQ865150) i *M. trossulus* (Mtp63/73-like, DQ865152). Razlike u nukleotidnim sljedovima istaknute su različitim bojama. Uokvireno su označeni očuvani motivi i regije vezanja regulatornih proteina; transkripcijski faktor Pax-4, transkripcijski faktor IRF-8, tumor-supresorski protein ING4, transkripcijski faktor Ets, transkripcijski faktor MEF-2C, proteini porodice Mushasi (MBE), K-box motiv (KB), poliadenilacijski signal (PAS).

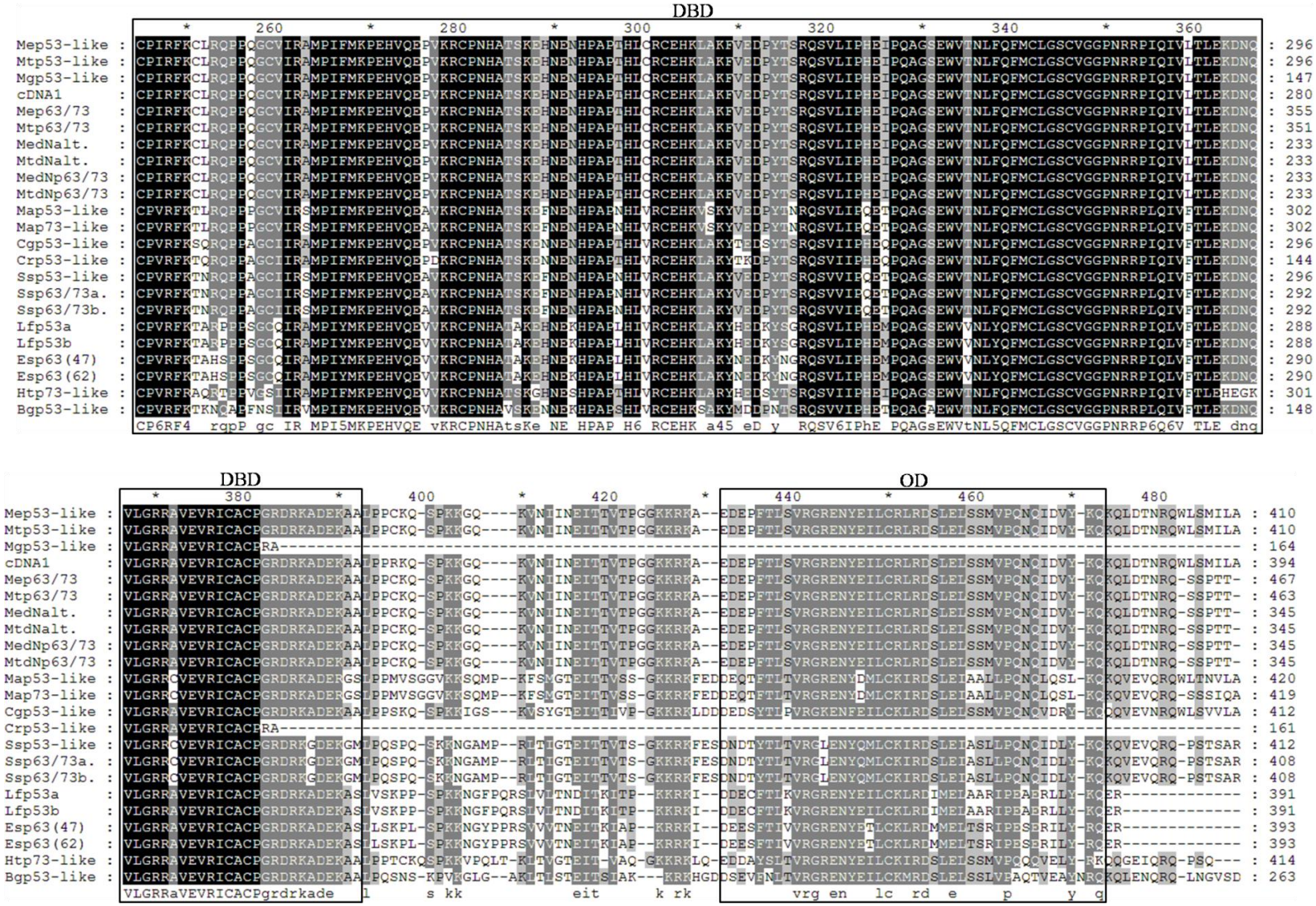
4.2. Usporedba kodirajućih sljedova porodice p53 mekušaca

Analiza svih pronađenih nukleotidnih sljedova porodice p53 mekušaca u GenBank bazi i dobivenih kodirajućih sljedova *M. galloprovincialis* (cDNA1-4) višestrukim sravnjenjima utvrdila je da svi sljedovi koji potječu od iste vrste imaju identičnu ili gotovo identičnu (razlika u svega nekoliko nukleotida na 1kb slijeda) tu središnju regiju, a ona kodira za DNA-veznu domenu (Slika 11). Sljedovi se najviše razlikuju (različita dužina, slabija očuvanost) u 5' i 3' krajevima. Izvedeni proteinski sljedovi svih pronađenih nukleotidnih sljedova porodice p53 mekušaca se mogu, prema sravnjenju na Slici 8 i usporedbom sa referentnim sljedovima porodice p53 čovjeka (sve izoforme gena u GenBank bazi) podijeliti u tri regije: regija N-kraja sa najmanje dvije alternativne forme (duža i kraća), središnja regija očuvana kod svih sljedova i regija C-kraja također sa najmanje dvije alternativne forme (duža i kraća). Središnja regija je kod svih mekušaca dužine oko 330 ak i u skladu sa odgovarajućim regijama proteina p63 i p73 čovjeka (sve izoforme) koji su dužine oko 340 ak. Cjelokupni protein p53 čovjeka (najduža izoforma, NP_000537) dužine je 393 ak. Uz DNA-veznu, središnja regija sadrži i oligomerizacijsku domenu kod svih analiziranih sljedova mekušaca te obje domene imaju visok stupanj očuvanosti. Iznimka su samo sljedovi „p53-like“ *M. galloprovincialis* (DQ158079) i *C. rhizophorae* (AY442309) kod kojih je cjelokupan izvedeni proteinski slijed samo regija DNA-vezne domene. Kraća regija N-kraja određena je sa 13 aminokiselina koje pokazuju visoki stupanj sličnosti (oko 80% aminokiselina dijeli ista fizikalno-kemijska svojstva na očuvanim mjestima) sa izvedenim proteinskim sljedovima egzona 3’ (ΔN fragment) gena p63 (GeneID 8626) i p73 (GeneID 7161) čovjeka (Slika 12). Samo sljedovi ΔNp63/p73 *M. edulis* i *M. trossulus* (DQ865151, DQ865153, DQ060436 i DQ060438) imaju identičnu kratku regiju 5’ kraja koja kodira za tih 13 očuvanih aminokiselina kraće regije N-kraja. Dok je kraća regija N-kraja kod svih izvedenih proteinskih sljedova jednake dužine (13 ak), duža regija N-kraja je kod većine oko 70 ak, ali kod *Mytilus* sljedova (DQ865150 i DQ865152) je 135 ak. Duža regija N-kraja izoformi p63 čovjeka je 108 ak, a izoformi p73 čovjeka 62 ak. Kod svih analiziranih sljedova mekušaca, duža regija N-kraja sadrži TA domenu porodice p53. Kraća regija C-kraja izvedenih proteinskih sljedova može se usporediti sa izvedenim proteinskim slijedom egzona 10’ koji je eksprimiran u γ izoformama proteina p63 čovjeka (Slika 13). Očuvanost je manja nego kod fragmenata ΔN, ali sličnost je ipak vidljiva jer značajan udio (oko 60%) aminokiselina dijeli ista fizikalno-kemijska svojstva na

očuvanim mjestima. Kodirajući sljedovi sa kraćom regijom C-kraja izvedenog proteina su: *Mytilus edulis* „p53-like“ (Mep53-like, AY579472), *Mytilus trossulus* „p53-like“ (Mtp53-like, AY611471), *Crassostrea gigas* „p53-like“ (Cgp53-like, AM236465), *Mya arenaria* „p53-like“ (Map53-like, AF253323), *Biomphalaria glabrata* „p53-like“ (Bgp53-like, GU929337) i *Mytilus galloprovincialis* cDNA1. Dužina kraćih regija C-kraja svih izvedenih proteina mekušaca je oko 30 aminokiselina što je u skladu sa dužinom (37 ak) izvedenog proteinskog slijeda egzona 10’ gena p63 čovjeka. Predstavnici roda *Mytilus* (*M. edulis* „p53-like“ (Mep53-like, AY579472), *M. trossulus* „p53-like“ (Mtp53-like, AY611471) i *M. galloprovincialis* cDNA1) imaju identične aminokiselinske sljedove tih kraćih C-krajeva. Duže regije C-kraja izvedenih proteina su oko 200 ak kod svih mekušaca, a kod p63 i p73 sljedova (α izoforme) čovjeka su nešto duže, oko 230 ak. Kod svih analiziranih sljedova mekušaca, duža regija C-kraja sadrži SAM domenu porodice p53 sa visokim stupnjem očuvanosti. Kodirajući sljedovi sa dužom regijom C-kraja izvedenog proteina su: *M. edulis* „p63/76-like“ (Mep63/73-l, DQ865150), *M. trossulus* „p63/76-like“ (Mtp63/73-l, DQ865152), *M. edulis* „ΔNp63/73-like“ (MedNp63/73, DQ865151, DQ040436), *M. trossulus* „ΔNp63/73-like“ (MtdNp63/73, DQ865153, DQ060438), *Loligo forbesi* p53a (Lfp53a, U43595), *Euprymna scolopes* p63(62) (Esp63(62), DQ247974), *Mya arenaria* „p73-like“ (Map73-like, AF253324), *Spisula solidissima* „p53-like“ p120 (Ssp53-like, AF285104), *Spisula solidissima* p63/73 alpha (Ssp63/73a., AY289767) i *Spisula solidissima* p63/73 beta (Ssp63/73b., AY289768)).

Jedini analizirani slijed mekušaca koji nije u skladu sa dobivenim rezultatima usporedbe sljedova porodice p53 je slijed p63(47) *Euprymna scolopes* (DQ247973) kod kojeg je izvedenih 37 aminokiselina kraćeg 3' kraja identično početku izvedenog proteinskog slijeda dužeg 3' kraja slijeda p63(62) *Euprymna scolopes* (DQ247974), a ne pokazuje sličnost sa izvedenim proteinskim slijedom egzona 10' kao svi ostali kraći 3' sljedovi. Osim toga, BLASTn analiza utvrdila je da slijed p53 *M. edulis* (AY705932) ima vrlo visoku sličnost (preko 80% identičnosti) sa nukleotidnim sljedovima porodice p53 riba pa nije korišten u ovome radu.

TA
DBD
Mep53-like
Mtp53-like
Mgp53-like
cDNA1
Mep63/73
Mtp63/73
MedNalt.
MtdNalt.
MedNp63/73
MtdNp63/73
Map53-like
Map73-like
Cgp53-like
Crp53-like
Ssp53-like
Ssp63/73a.
Ssp63/73b.
Lfp53a
Lfp53b
Esp63(47)
Esp63(62)
Htp73-like
Bgp53-like


DBD
260
280
300
320
340
360
Mep53-like
Mtp53-like
Mgp53-like
cDNA1
Mep63/73
Mtp63/73
MedNalt.
MtdNalt.
MedNp63/73
MtdNp63/73
Map53-like
Map73-like
Cgp53-like
Crp53-like
Ssp53-like
Ssp63/73a.
Ssp63/73b.
Lfp53a
Lfp53b
Esp63(47)
Esp63(62)
Htp73-like
Bgp53-like
DBD
OD
380
400
420
440
460
480

```
                                                                                                        SAM
               *        500         *        520         *        540         *        560         *        580         *        600         *
Mep53-like  : RENKNKLMKKVKRPQHRPGIKSRT-------------------------------------------------------------------------------------------------- : 434
Mtp53-like  : RENKNKLMKKVKRPQHRPGIKSRT-------------------------------------------------------------------------------------------------- : 434
Mgp53-like  : -------------------------------------------------------------------------------------------------------------------------- :   -
cDNA1       : RENKNKLMKKVKRPQHRPGIKSRT-------------------------------------------------------------------------------------------------- : 418
Mep63/73    : STARVVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQNGVVENHGN---------IKEE----LMANGDHSISNWLTTLGLSAYIDNFHQQNLFTMEQLDDFTVEDLQKMRIGTSHRN : 576
Mtp63/73    : --ARVVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQNGVVENHGN---------IKEE----LMANGDHSISNWLTTLGLSAYIDNFHQQNLFTMEQLDDFTVEDLQKMRIGTSHRN : 570
MedNalt.    : STARVVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQNGVVENHGN---------IKEE----LMANGDHSISNWLTTLGLSAYIDNFHQQNLFTMEQLDDFTVEDLQKMRIGTSHRN : 454
MtdNalt.    : --ARVVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQNGVVENHGN---------IKEE----LMANGDHSISNWLTTLGLSAYIDNFHQQNLFTMEQLDDFTVEDLQKMRIGTSHRN : 452
MedNp63/73  : STARVVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQNGVVENHGN---------IKEE----LMANGDHSISNWLTTLGLSAYIDNFHQQNLFTMEQLDDFTVEDLQKMRIGTSHRN : 454
MtdNp63/73  : --ARVVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQNGVVENHGN---------IKEE----LMANGDHSISNWLTTLGLSAYIDNFHQQNLFTMEQLDDFTVEDLQKMRIGTSHRN : 452
Map53-like  : KEGKSRLIKKKHRPGKIIRHPLK--------------------------------------------------------------------------------------------------- : 443
Map73-like  : ATSSARIPAIAATPATTYVQQGVTTTSSDGKQLTMPFNTQELVQVTSSDVSHDGAVPQPIKEE--TIQNDMQDNSVSTWLNALGLGAYIDGFHEQNLYSLLQLDDFSLDDLAKMKIGNSHRN : 539
Cgp53-like  : RESKNKLMKRKTKPGKVIKRPA---------------------------------------------------------------------------------------------------- : 434
Crp53-like  : -------------------------------------------------------------------------------------------------------------------------- :   -
Ssp53-like  : VLAVPASQAGQVVPQTTPELRQDTQPFVSQVE-----------------------VPQPIKEETITENGEMTDISVAAWLNHLGLGAYIDSFHEHNLYSVIQLDDFSLDDLAKMKIGNAHRN : 511
Ssp63/73a.  : VLAVPASQAGQVVPQTTPELRQDTQPFVSQVE-----------------------VPQPIKEETITENGEMTDISVAAWLNHLGLGAYIDSFHEHNLFSVIQLDDFSLDDLAKMKIGNAHRN : 507
Ssp63/73b.  : VLAVPASQAGQVVPQTTPELRQDTQPFVSQVE-----------------------VPQPIKEETITENGEMTDISVAAWLNHLGLGAYIDSFHEHNLYSVIQLDDFSLDDLAKMKIGNAHRN : 507
Lfp53a      : -------------------------------------------------------------------------------------------------------------------------- :   -
Lfp53b      : QAPIG---RLTSLPSSSSNGSQDGSRSSTAFSTSDSSQ---VNSSQNNTQ-MVNGQ-VPHEEETPVTKCEPTENTIAQWLTKLGLQAYIDNFQQKGLHNMFQLDEFTLEDLQSMRIGTGHRN : 505
Esp63(47)   : QAMLGHLPSTSSNGSQDGSRSSSTSYSTSDKSQVNSSQ------------------------------------------------------------------------------------ : 431
Esp63(62)   : QAMLGHLPSTSSNGSQDGSRSSSTSYSTSDKSQVNSSQ--------NNSQRLVNGQQVTHEEETQVPKCEPNENTIAQWLTKLGLQAYIDNFQQKGLHNMFQLDEFTLEDLQSMRIGTGHRN : 507
Htp73-like  : -PHVISIPSNERIAQPAETPHQASLPFSPDSSNVSSSQNNQLNGDTLNAAMEAVINQAAVKEEVVPNGQTLVDNSVASWLTSIGLSAYIDNFHEKNFLNMFQLDEFSLEDLSAMKIGTSHRN : 535
Bgp53-like  : YGSDNSEASTSSTMSIEPHHPSLGSYEVSDALNDLFDTVTSSHSPQHDSD------LAVKKEDIDVALGDLRDNSVAGWLDRLGLSNYIDRFHQNGYDNLFQLDECKKETLEHLGIPEVDRN : 379
```

```
                  SAM
                  620         *        640         *        660         *        680         *        700
Mep53-like  : ----------------------------------------------------------------------------------------------- :   -
Mtp53-like  : ----------------------------------------------------------------------------------------------- :   -
Mgp53-like  : ----------------------------------------------------------------------------------------------- :   -
cDNA1       : ----------------------------------------------------------------------------------------------- :   -
Mep63/73    : KIWKALVEFHSESITIS--DSQSLQRDVSTA-STISMTSQASISQNS--QNSTYCPGYYEVTRYTFKHTVSYTKVDER--SPKRA-KVD------ : 657
Mtp63/73    : KIWKALVEFHSESITIS--DSQSLQRDVSTA-STISMTSQASISQNS--QNSTYCPGYYEVTRYTFKHTVSYTKVDER--SPKRA-KVD------ : 651
MedNalt.    : KIWKALVEFHSESITIS--DSQSLQRDVSTA-STISMTSQASISQNS--QNSTYCPGYYEVTRYTFKHTVSYTKVDER--SPKRA-KVD------ : 535
MtdNalt.    : KIWKALVEFHSESITIS--DSQSLQRDVSTA-STISMTSQASISQNS--QNSTYCPGYYEVTRYTFKHTVSYTKVDER--SPKRA-KVD------ : 533
MedNp63/73  : KIWKALVEFHSESITIS--DSQSLQRDVSTA-STISMTSQASISQNS--QNSTYCPGYYEVTRYTFKHTVSYTKVDER--SPKRA-KVD------ : 535
MtdNp63/73  : KIWKALVEFHSESITIS--DSQSLQRDVSTA-STISMTSQASISQNS--QNSTYCPGYYEVTRYTFKHTVSYTKVDER--SPKRA-KVD------ : 533
Map53-like  : ----------------------------------------------------------------------------------------------- :   -
Map73-like  : KIWKSLLELRNQGFTTAESQDSLA--KQASSTSTISLASQGSISQ-----NSTYNPGFYEVTRHTFKHTISLTKEERHVQTRSSASKAD------ : 621
Cgp53-like  : ----------------------------------------------------------------------------------------------- :   -
Crp53-like  : ----------------------------------------------------------------------------------------------- :   -
Ssp53-like  : KIWKSVLELRNEGLTAIQADSQDLQ-KNASSTSTLSITSQ-----NSVSQNSTYNPGYYEVTRYTFKHTISLTAEEKKLQELSPPT--------- : 591
Ssp63/73a.  : KIWKSVLELRNEGLTAIQADSQDLQ-KNASSTSTLSITSQ-----NSVSQNSTYNPGYYEVTRYTFKHTISLTAEEKKLQELSPPTKKEKLSTDV : 596
Ssp63/73b.  : KIWKSVLELRNEGLTAIQADSQDLQ-KNASSTSTLSITSQ-----NSVSQNSTYNPGYYEVTRYTFKHTISLTAEEKKLQELSPPTKKEKLPTDV : 596
Lfp53a      : ----------------------------------------------------------------------------------------------- :   -
Lfp53b      : KIWKSLLDYRRLLSSGTESQALQHAASNASTLSVGSQNSYCPGFYEVTRYTYKHTISYL------------------------------------ : 564
Esp63(47)   : ----------------------------------------------------------------------------------------------- :   -
Esp63(62)   : KIWKSLLDYRRVLSTGTETQALQHAASNASTLSVGSQNSYCPGYYEVTRYSFKHTISYL------------------------------------ : 566
Htp73-like  : KIWKSLVEYKQANAYTSEVSQSLA--RNASSTSTISLSSQPSISQNS--AYCPGYYEVTRYTFKHTISLT--KDEAKKLKTEKPE---------- : 614
Bgp53-like  : KIWKSLVEFHQYN--QIQDMTQAFVGRSNSRSSTESTTTLVFPTQSSATSS-------------------------------------------- : 428
```

Slika 11. Višestruko sravnjenje izvedenih proteinskih sljedova cDNA1 *M.galloprovincialis* i ostalih kodirajućih sljedova porodice p53 mekušaca, dostupnih u GenBank bazi podataka: *Mytilus edulis* „p53-like“ (Mep53-like, AY579472), *Mytilus trossulus* „p53-like“ (Mtp53-like, AY611471), *Mytilus galloprovincialis* „p53-like“ (Mgp53-like, DQ158079), *M. edulis* „p63/76-like“ (Mep63/73, DQ865150), *M. trossulus* „p63/76-like“ (Mtp63/73, DQ865152), *M. edulis* „ΔNp63/73-like“ (MedNp63/73, DQ865151), *M. edulis* „ΔNp63/73-like, alternatively spliced“ (MedNalt., DQ060436), *M. trossulus* „ΔNp63/73-like“ (MtdNp63/73, DQ865153), *M. trossulus* „ΔNp63/73-like, alternatively spliced“ (MtdNalt., DQ060438), *Mya arenaria* „p53-like“ (Map53-like, AF253323), *Mya arenaria* „p73-like“ (Map73-like, AF253324), *Crassostrea gigas* „p53-like“ (Cgp53-like, AM236465), *Crassostrea rhizophorae* „p53-like“ (Crp53-like, AY442309), *Spisula solidissima* „p53-like“ p120 (Ssp53-like, AF285104), *Spisula solidissima* p63/73alpha (Ssp63/73a., AY289767), *Spisula solidissima* p63/73beta (Ssp63/73b., AY289768), *Haliotis tuberculata* „p73-like“ (Htp73-like, AM396936), *Loligo forbesi* p53a (Lfp53a, U43595), *Loligo forbesi* p53b (Lfp53b, U43596), *Euprymna scolopes* p63(47) (Esp63(47), DQ247973), *Euprymna scolopes* p63(62) (Esp63(62), DQ247974). Crno uokvirene regije označavaju očuvane domene porodice p53. Crveno uokvireno je označena kraća alternativna regija N-kraja (ΔN fragment). Sravnjenje je dodatno ručno podešeno kako bi se uskladile regije očuvanih domena porodice p53 svih sljedova.

```
Hsp73dN   : MLYVGDPAR-HLAT : 13
Hsp63dN   : MLYLENNAQTQFSE : 14
MytilusdN : MIKFERTGFTTYR- : 13
```

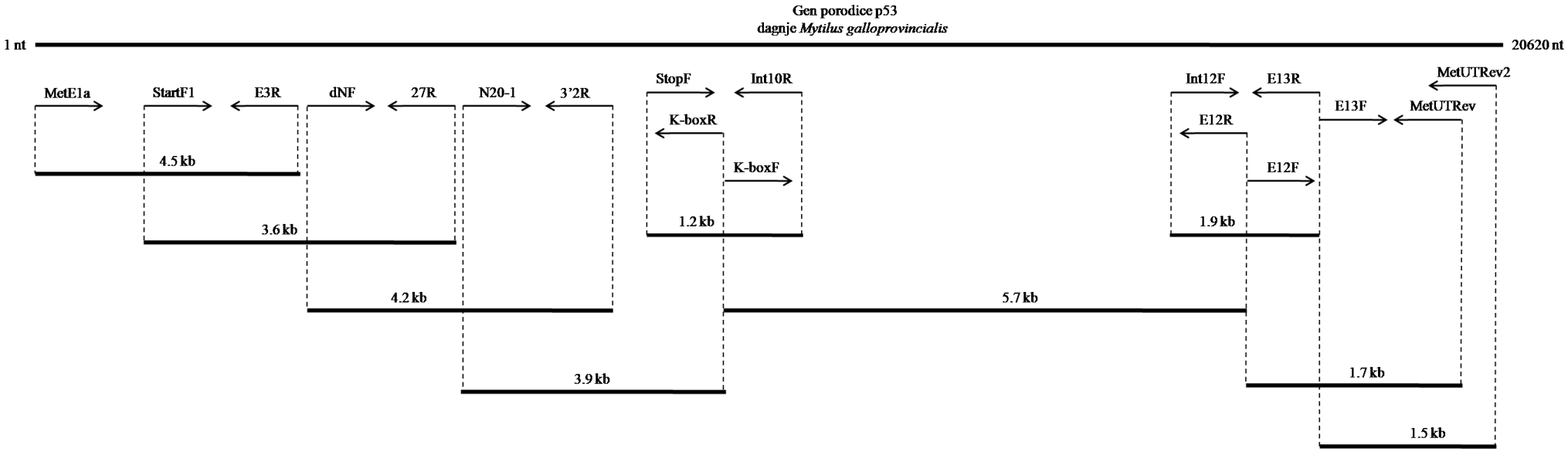
Slika 12. Višestruko sravnjenje izvedenih proteinskih sljedova egzona 3' gena p63 i p73 čovjeka (Hsp73dN, GeneID 7161 i Hsp63dN, GeneID 8626) i 5' kraćih krajeva *Mytilus* „ΔNp63/73-like“ sljedova (MytilusdN, DQ865151, DQ040436, DQ865153 i DQ060438) koji su 100% identični u ovoj regiji. Bojanjem (prema postavkama računalnog programa GeneDoc) istaknuta su fizikalno-kemijska svojstva aminokiselina.

```
Map53l     : -----WLTNVLA--KEG--KSRLIKKKHRPGK--IIRHPLK : 30
Cgp53l     : -----WLSVVLA--RES--KNKLMKRKTKPGK--VIKRPA- : 29
Mytilus    : -----WLSMILA--REN--KNKLMKKVKRPQHRPGIKSRI- : 31
Hsp63_10'  : LLSACFRNELVEFRRETEKQSDVFFRHSKPPN--RSVYP-- : 37
```

Slika 13. Višestruko sravnjenje izvedenih proteinskih sljedova egzona 10' gena p63 čovjeka (Hsp63_10', GeneID 8626) i 3' kraćeg kraja kodirajućih DNA porodice p53 mekušaca (*Mya arenaria* (Map53l, AF253323), *Crassostrea gigas* (Cgp53l, AM236465) i u ovoj regiji 100% identičnih sljedova roda *Mytilus* (Mytilus, AY579472, AY611471 i cDNA1). Bojanjem (prema postavkama računalnog programa GeneDoc) istaknuta su fizikalno-kemijska svojstva aminokiselina.

4.3. Genomski slijed gena porodice p53 *Mytilus galloprovincialis*

Cjeloviti genomski slijed gena porodice p53 dobiven je kao konsenzus slijed svih sekvenciranih fragmenata genomske DNA (Slika 14) i ima prosječnu pokrivenost od 3,75. Genomski slijed pohranjen je u NCBI GenBank bazu podataka pod pristupnim brojem KP233801. Genomski slijed sadrži 20620 nukleotida (nt). Višestrukim sravnjenjima (Lasergene 7) genomskog slijeda i kodirajućih sljedova dobivenih iz cDNA biblioteke (cDNA1-4) određena je egzon-intron struktura gena porodice p53 *M. galloprovincialis*. Za jasniji prikaz dobivenih rezultata sravnjenja korišten je Splign program dostupan na mreži (Slika 15). Uočeno je ukupno 13 različitih nukleotida (eng. *single nucleotide polymorphism*, SNP) između genomske i kodirajućih DNA sljedova (cDNA1-4). Od 14 razlika, 9 je tihih (na genomskoj DNA; nt 4830 (T/C), 8778 (A/T), 9786 (C/T), 9882 (A/G), 10168 (C/A), 10174 (A/T), 10202 (C/A), 10764 (T/A) i 16600 (G/T)) te ne mijenjaju pripadajuću aminokiselinu

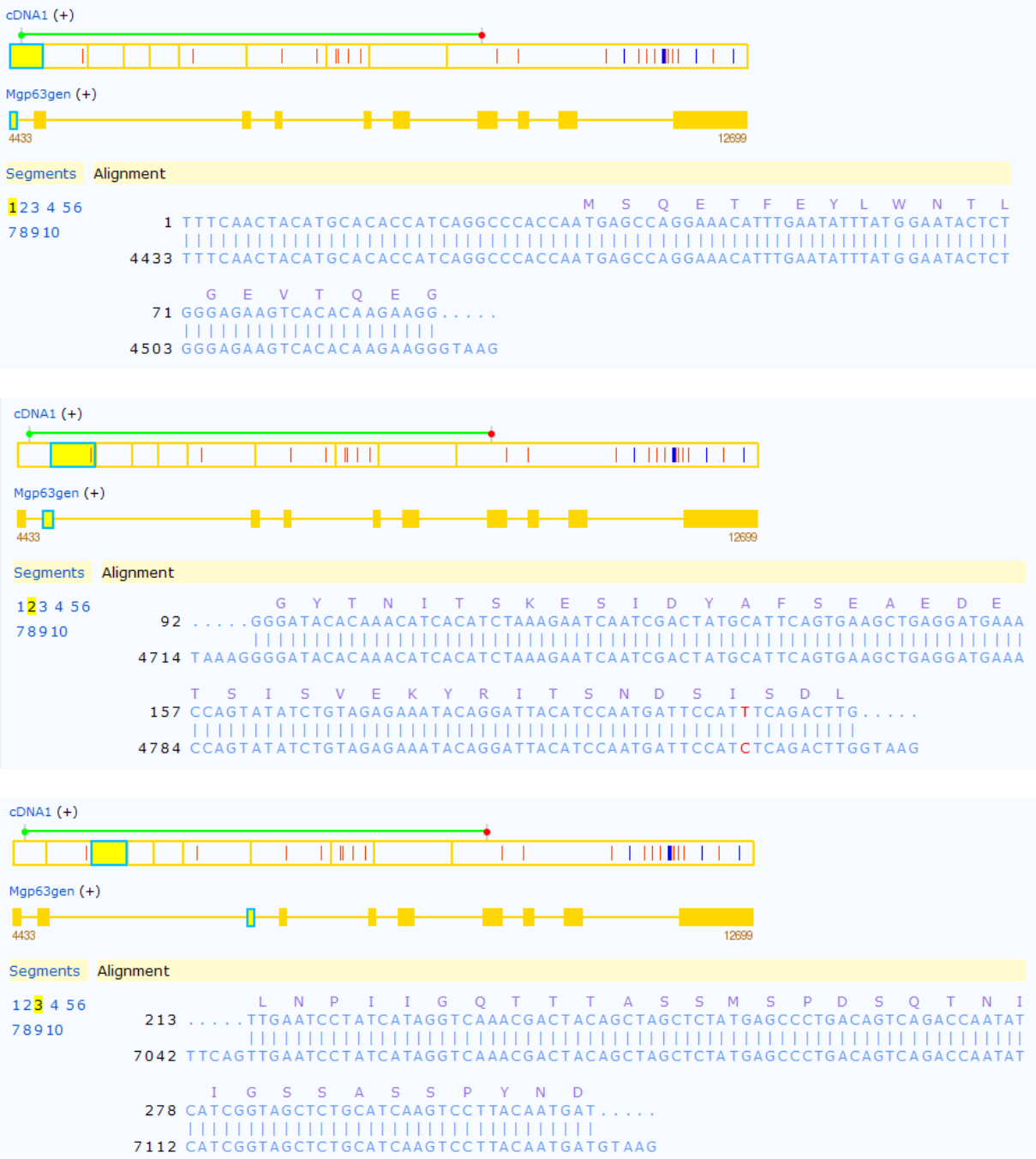


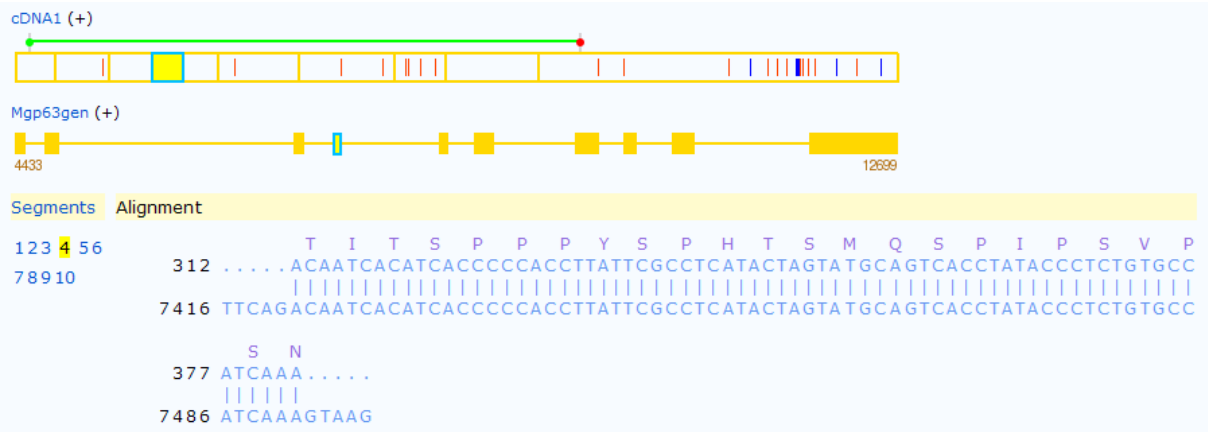
Slika 14. Shematski prikaz strategije umnažanja i sekvenciranja fragmenata gena porodice p53 dagnje *Mytilus galloprovincialis*. Umnoženi fragmenti gena porodice p53 naznačeni podebljanim linijama, a strelice označavaju sljedove korištenih početnica.

u izvedenom proteinskom slijedu gena porodice p53 i većinom znače promjenu treće baze kodona. Pet razlika (na genomskoj DNA; nt 10235 (C/T), 10237 (C/T), 10786 (T/A), 16582(C/G) i 18397(A/T)) mijenja aminokiselinu u izvedenom proteinskom slijedu (Slika 15). Prema sravnjenjima, genomski slijed *M. galloprovincialis* kodira gen sastavljen od 14 egzona isprekidanih sa 13 introna. Slijed cDNA1 otkrio je 10 egzona. Prema izvedenom proteinskom slijedu utvrđena je sličnost zadnjeg u nizu egzona sa egzonom 10' kojim je određena γ izoforma C-kraja proteina p63 čovjeka (Slika 13). Izvedeni proteinski slijed cDNA1 potvrdio je i očuvane domene porodice p53 (TA-prvi egzon u nizu, DBD-dio petog i osmog egzona te cjeloviti egzoni 6 i 7 i OD- deveti egzon u nizu) pa je niz od deset egzona koji kodiraju cDNA1 numeriran u strukturi gena porodice p53 kao egzoni 2 do 10 te još zadnji egzon 10'. Slijed cDNA2 potvrdio je egzone 10 i 10' te utvrdio ukupnu dužinu UTR-a egzona 10' od 887 nt. Ovaj unutrašnji UTR gena odgovara nukleotidnoj regiji 11981-12867 nt genomske DNA. Uočeno je 15 razlika između genomske i kodirajućih DNA sljedova u regiji unutrašnjeg UTR-a. Ove razlike nisu samo polimorfizmi jednog nukleotida (kao što je utvrđeno u translatiranim regijama egzona) nego i delecije/insercije jednog ili više nukleotida. Početnih 77 nukleotida cDNA3 odgovara (100% identičnost) regiji 3549-3625 nt genomskog slijeda te predstavlja prvi egzon u strukturi gena porodice p53 *M. galloprovincialis.* Sravnjenja genomskog slijeda sa reprezentativnim poznatim kodirajućim sljedovima porodice p53 *M. edulis* i *M. trossulus* iz GenBank baze podataka („p63/73-like“ DQ865150 i DQ865152 te „p53-like“;, DQ060435 i DQ060437) potvrdila su utvrđenu strukturu gena porodice p53 *M.galloprovincialis* (14 egzona i 13 introna), ali „p63/73-like“ sljedovi kao prvi egzon gena porodice p53 pretpostavljaju regiju 1-106 nt genomske DNA *M. galloprovincialis* koja odgovara (100% identičnost) regiji 41-146 nt „p63/73-like“ kodirajućih sljedova *M. edulis* (DQ865150) i *M. trossulus* (DQ865152) (Slika 16). Kodirajući sljedovi „p63/p73-like“ *M. edulis* i *M. trossulus* imaju tako i produženje od 40 nukleotida na 5' kraju u odnosu na genomsku DNA *M. galloprovincialis*. Kodirajući sljedovi „p53-like“ *M. edulis* i *M. trossulus* započinju u egzonu 2 (pozicija nt 4412 genomskog slijeda *M. galloprovincialis*) te određuju početak translacije (ATG) na poziciji 4443-4445 nt genomskog slijeda. Prvi egzon gena porodice p53 *M.galloprovincialis* utvrđen slijedom cDNA3 je označen kao egzon 1', a prvi egzon utvrđen sljedovima „p63/73-like“ *M. edulis* i *M. trossulus* kao egzon 1. U skladu sa izvedenim proteinskim slijedom egzona 2 i poštujući zadani okvir čitanja, egzon 1 može biti kodirajući dok je egzon 1' sigurno nekodirajući (Slika 16). Mogući početak translacije (ATG)

unutar egzona 1, sa odgovarajućim konsenzus Kozak slijedom (Kozak, 1987), nalazi se na poziciji nt 75-77 genomskog slijeda.

(A)

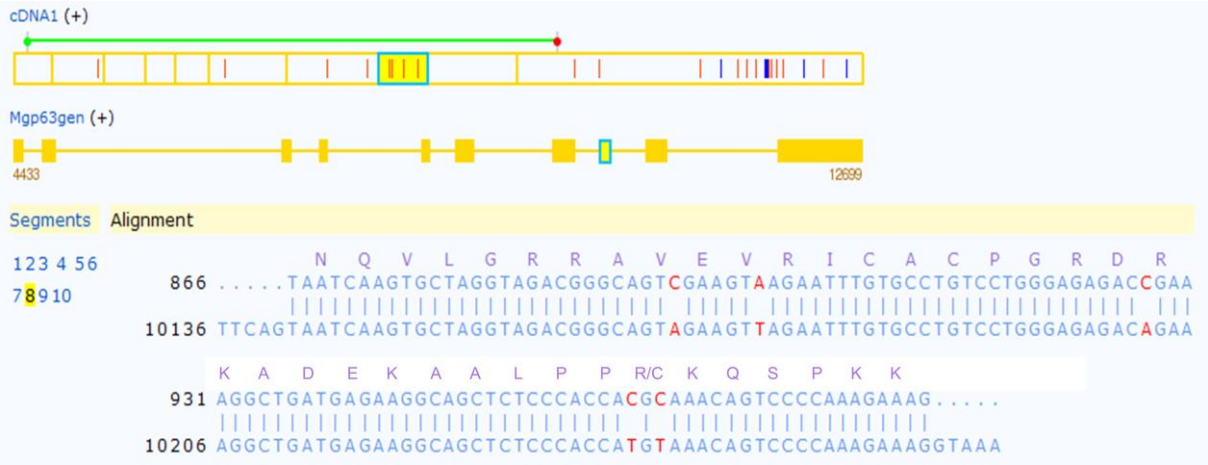


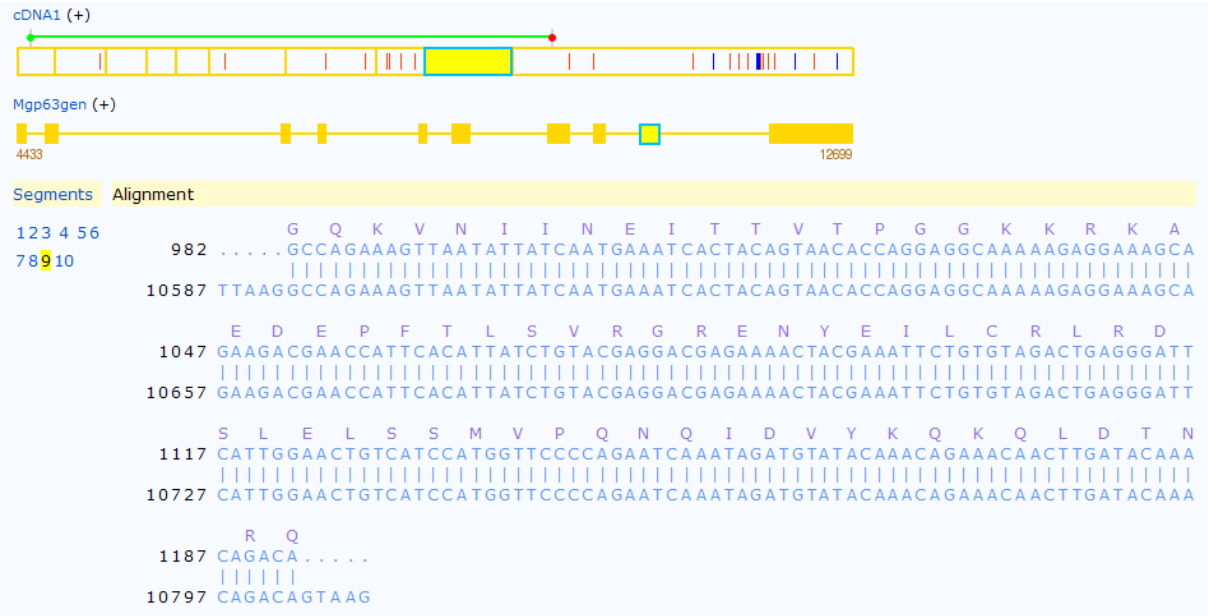

cDNA1 (+)
Mgp63gen (+)
4433
12699
Segments
Alignment
123 4 56
78910
T I T S P P P Y S P H T S M Q S P I P S V P
312 .....ACAATCACATCACCCCCACCTTATTCGCCTCATACTAGTATGCAGTCACCTATACCCTCTGTGCC
7416 TTCAGACAATCACATCACCCCCACCTTATTCGCCTCATACTAGTATGCAGTCACCTATACCCTCTGTGCC
S N
377 ATCAAA.....
7486 ATCAAAGTAAG

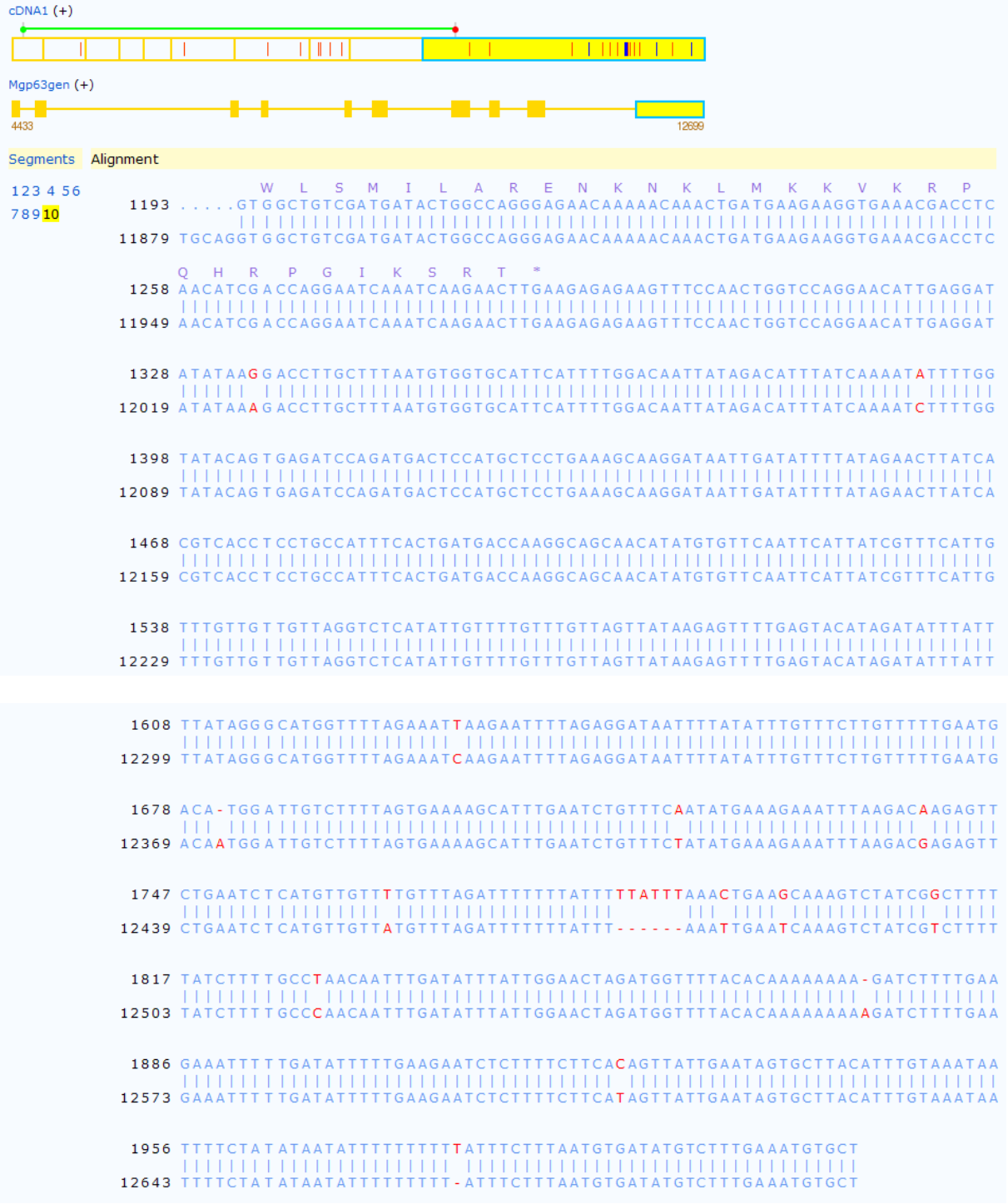
cDNA1 (+)
Mgp63gen (+)
4433
12699
Segments
Alignment
123 4 56
78910
T D Y P G D Y G F T I S F S Q P S K E T K
383 .....CACAGACTATCCCGGCGATTATGGGTTCACCATTTCATTTTCTCAGCCATCAAAAGAGACAAAGT
8410 TACAGCACAGACTATCCCGGCGATTATGGGTTCACCATTTCATTTTCTCAGCCATCAAAAGAGACAAAGT
S T T W T
448 CTACCACATGGACA.....
8480 CTACCACATGGACAGTAAG

cDNA1 (+)
Mgp63gen (+)
4433
12699
Segments
Alignment
123 4 56
78910
Y S E S L K K L Y V R M A T T C P I R F K C
462 .....TACTCAGAGTCACTTAAGAAATTATACGTCAGAATGGCAACAACTTGCCCAATCCGGTTTAAGTG
8732 TTCAGTACTCAGAGTCACTTAAGAAATTATACGTCAGAATGGCAACTACTTGCCCAATCCGGTTTAAGTG
L R Q P P Q G C V I R A M P I F M K P E H V Q
527 TTTGAGACAGCCTCCACAAGGATGTGTTATTCGTGCAATGCCAATATTCATGAAACCCGAACATGTCCAA
8802 TTTGAGACAGCCTCCACAAGGATGTGTTATTCGTGCAATGCCAATATTCATGAAACCCGAACATGTCCAA
E P V K R C P N H A T S K E H N E
597 GAACCTGTAAAAAGATGCCCAAATCATGCCACATCTAAAGAGCACAATGAAA.....
8872 GAACCTGTAAAAAGATGCCCAAATCATGCCACATCTAAAGAGCACAATGAAAGTAAG

cDNA1 (+)
Mgp63gen (+)
4433
12699
Segments
Alignment
123 4 56
7 8 9 10
N H P A P T H L C R C E H K L A K F V E D P
649 .....ATCATCCAGCTCCAACACATTTATGTCGATGTGAGCACAAACTTGCTAAATTTGTTGAAGATCCA
9684 TTTAGATCATCCAGCTCCAACACATTTATGTCGATGTGAGCACAAACTTGCTAAATTTGTTGAAGATCCA
Y T S R Q S V L I P H E I P Q A G S E W V T N
714 TATACCAGCCGCCAGAGTGTTCTAATTCCACACGAGATACCTCAAGCTGGCTCAGAATGGGTCACCAATT
9754 TATACCAGCCGCCAGAGTGTTCTAATTCCACATGAGATACCTCAAGCTGGCTCAGAATGGGTCACCAATT
L F Q F M C L G S C V G G P N R R P I Q I V L T
784 TGTTCCAGTTCATGTGCCTGGGGTCATGTGTAGGAGGACCAAACAGAAGGCCTATTCAAATTGTTCTGAC
9824 TGTTCCAGTTCATGTGCCTGGGGTCATGTGTAGGAGGACCAAACAGAAGGCCTATTCAGATTGTTCTGAC
L E K D
854 TTTAGAAAAAGA.....
9894 TTTAGAAAAAGAGTAAG


cDNA1 (+)
Mgp63gen (+)
4433
12699
Segments
Alignment
123 4 56
7 8 9 10
N Q V L G R R A V E V R I C A C P G R D R
866 .....TAATCAAGTGCTAGGTAGACGGGCAGTCGAAGTAAGAATTTGTGCCTGTCCTGGGAGAGACCGAA
10136 TTCAGTAATCAAGTGCTAGGTAGACGGGCAGTAGAAGTTAGAATTTGTGCCTGTCCTGGGAGAGACAGAA
K A D E K A A L P P R/C K Q S P K K
931 AGGCTGATGAGAAGGCAGCTCTCCCACCACGCAAACAGTCCCCAAAGAAAG.....
10206 AGGCTGATGAGAAGGCAGCTCTCCCACCATGTAAACAGTCCCCAAAGAAAGGTAAA


cDNA1 (+)
Mgp63gen (+)
4433
12699
Segments
Alignment
123 4 56
7 8 9 10
G Q K V N I I N E I T T V T P G G K K R K A
982 .....GCCAGAAAGTTAATATTATCAATGAAATCACTACAGTAACACCAGGAGGCAAAAAGAGGAAAGCA
10587 TTAAGGCCAGAAAGTTAATATTATCAATGAAATCACTACAGTAACACCAGGAGGCAAAAAGAGGAAAGCA
E D E P F T L S V R G R E N Y E I L C R L R D
1047 GAAGACGAACCATTCACATTATCTGTACGAGGACGAGAAAACTACGAAATTCTGTGTAGACTGAGGGATT
10657 GAAGACGAACCATTCACATTATCTGTACGAGGACGAGAAAACTACGAAATTCTGTGTAGACTGAGGGATT
S L E L S S M V P Q N Q I D V Y K Q K Q L D T N
1117 CATTGGAACTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAACAGAAACAACTTGATACAAA
10727 CATTGGAACTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAACAGAAACAACTTGATACAAA
R Q
1187 CAGACA.....
10797 CAGACAGTAAG


cDNA1 (+)
Mgp63gen (+)
4433
12699
Segments
Alignment
123 4 56
789 10
W L S M I L A R E N K N K L M K K V K R P
1193 .....GTGGCTGTCGATGATACTGGCCAGGGAGAACAAAAACAAACTGATGAAGAAGGTGAAACGACCTC
11879 TGCAGGTGGCTGTCGATGATACTGGCCAGGGAGAACAAAAACAAACTGATGAAGAAGGTGAAACGACCTC
Q H R P G I K S R T *
1258 AACATCGACCAGGAATCAAATCAAGAACTTGAAGAGAGAAGTTTCCAACTGGTCCAGGAACATTGAGGAT
11949 AACATCGACCAGGAATCAAATCAAGAACTTGAAGAGAGAAGTTTCCAACTGGTCCAGGAACATTGAGGAT
1328 ATATAAGGACCTTGCTTTAATGTGGTGCATTCATTTTGGACAATTATAGACATTTATCAAAATATTTTGG
12019 ATATAAAGACCTTGCTTTAATGTGGTGCATTCATTTTGGACAATTATAGACATTTATCAAAATCTTTTGG
1398 TATACAGTGAGATCCAGATGACTCCATGCTCCTGAAAGCAAGGATAATTGATATTTTATAGAACTTATCA
12089 TATACAGTGAGATCCAGATGACTCCATGCTCCTGAAAGCAAGGATAATTGATATTTTATAGAACTTATCA
1468 CGTCACCTCCTGCCATTTCACTGATGACCAAGGCAGCAACATATGTGTTCAATTCATTATCGTTTCATTG
12159 CGTCACCTCCTGCCATTTCACTGATGACCAAGGCAGCAACATATGTGTTCAATTCATTATCGTTTCATTG
1538 TTTGTTGTTGTTAGGTCTCATATTGTTTTGTTTGTTAGTTATAAGAGTTTTGAGTACATAGATATTTATT
12229 TTTGTTGTTGTTAGGTCTCATATTGTTTTGTTTGTTAGTTATAAGAGTTTTGAGTACATAGATATTTATT
1608 TTATAGGGCATGGTTTTAGAAATTAAGAATTTTAGAGGATAATTTTATATTTGTTTCTTGTTTTTGAATG
12299 TTATAGGGCATGGTTTTAGAAATCAAGAATTTTAGAGGATAATTTTATATTTGTTTCTTGTTTTTGAATG
1678 ACA-TGGATTGTCTTTTAGTGAAAAGCATTTGAATCTGTTTCAATATGAAAGAAATTTAAGACAAGAGTT
12369 ACAATGGATTGTCTTTTAGTGAAAAGCATTTGAATCTGTTTCTATATGAAAGAAATTTAAGACGAGAGTT
1747 CTGAATCTCATGTTGTTTTGTTTAGATTTTTTTATTTTTATTTAAACTGAAGCAAAGTCTATCGGCTTTT
12439 CTGAATCTCATGTTGTTATGTTTAGATTTTTTTATTT------AAATTGAATCAAAGTCTATCGTCTTTT
1817 TATCTTTTGCCTAACAATTTGATATTTATTGGAACTAGATGGTTTTACACAAAAAAAA-GATCTTTTGAA
12503 TATCTTTTGCCCAACAATTTGATATTTATTGGAACTAGATGGTTTTACACAAAAAAAAAGATCTTTTGAA
1886 GAAATTTTTGATATTTTTGAAGAATCTCTTTTCTTCACAGTTATTGAATAGTGCTTACATTTGTAAATAA
12573 GAAATTTTTGATATTTTTGAAGAATCTCTTTTCTTCATAGTTATTGAATAGTGCTTACATTTGTAAATAA
1956 TTTTCTATATAATATTTTTTTTTTATTTCTTTAATGTGATATGTCTTTGAAATGTGCT
12643 TTTTCTATATAATATTTTTTTTT-ATTTCTTTAATGTGATATGTCTTTGAAATGTGCT

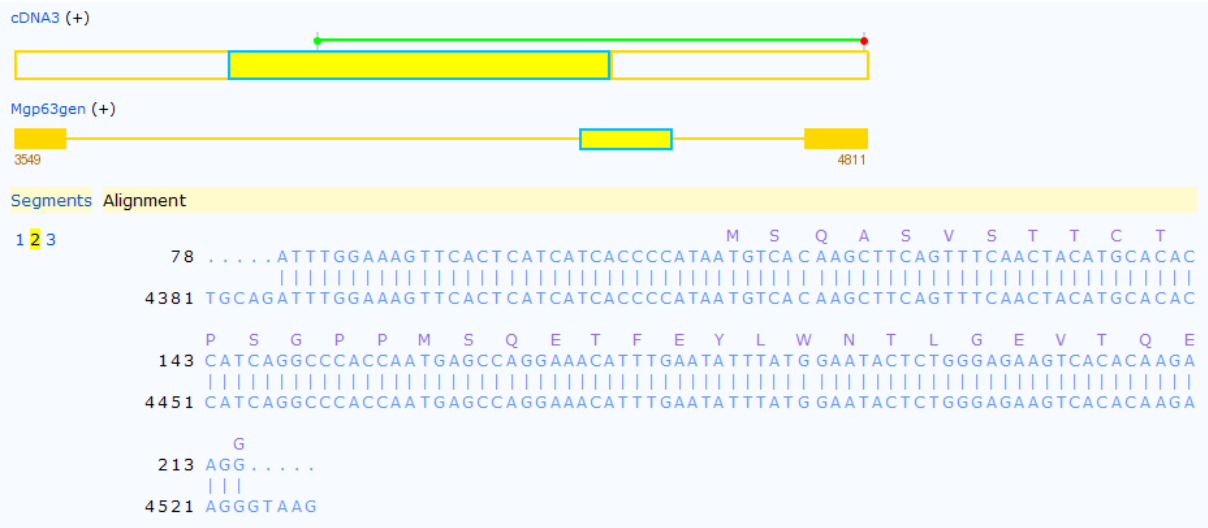
(B)

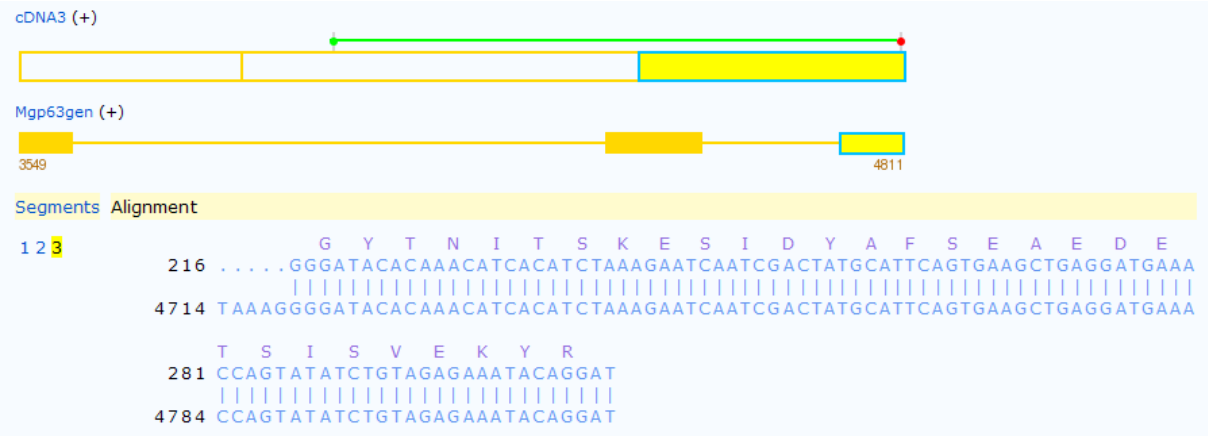
cDNA2 (+)
Mgp63gen (+)
10707
12863
Segments Alignment
1 2
L C R L R D S L E L S S M V P Q N Q I D V Y K
1 TCTGTGTAGACTGAGGGATTCATTGGAACTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAA
10707 TCTGTGTAGACTGAGGGATTCATTGGAACTGTCATCCATGGTTCCCCAGAATCAAATAGATGTATACAAA
Q K Q L D T N R Q
71 CAGAAACAACTTGATACAAACAGACA.....
10777 CAGAAACAACTTGATACAAACAGACAGTAAG

cDNA2 (+)
Mgp63gen (+)
10707
12863
Segments Alignment
1 2
W L S M I L A R E N K N K L M K K V K R P
97 .....GTGGCTGTCGATGATACTGGCCAGGGAGAACAAAAACAAACTGATGAAGAAGGTGAAACGACCTC
11879 TGCAGGTGGCTGTCGATGATACTGGCCAGGGAGAACAAAAACAAACTGATGAAGAAGGTGAAACGACCTC
Q H R P G I K S R T *
162 AACATCGACCAGGAATCAAATCAAGAACTTGAAGAGAGAAGTTTCCAACTGGTCCAGGAACATTGAGGAT
11949 AACATCGACCAGGAATCAAATCAAGAACTTGAAGAGAGAAGTTTCCAACTGGTCCAGGAACATTGAGGAT
232 ATATAAGGACCTTGCTTTAATGTGGTGCATTCATTTTGGACAATTATAGACATTTATCAAAATATTTTGG
12019 ATATAAAGACCTTGCTTTAATGTGGTGCATTCATTTTGGACAATTATAGACATTTATCAAAATCTTTTGG
302 TATACAGTGAGATCCAGATGACTCCATGCTCCTGAAAGCAAGGATAATTGATATTTTATAGAACTTATCA
12089 TATACAGTGAGATCCAGATGACTCCATGCTCCTGAAAGCAAGGATAATTGATATTTTATAGAACTTATCA
372 CGTCACCTCCTGCCATTTCACTGATGACCAAGGCAGCAACATATGTGTTCAATTCATTATCGTTTCATTG
12159 CGTCACCTCCTGCCATTTCACTGATGACCAAGGCAGCAACATATGTGTTCAATTCATTATCGTTTCATTG
442 TTTGTTGTTGTTAGGTCTCATATTGTTTTGTTTGTTAGTTATAAGAGTTTTGAGTACATAGATATTTATT
12229 TTTGTTGTTGTTAGGTCTCATATTGTTTTGTTTGTTAGTTATAAGAGTTTTGAGTACATAGATATTTATT

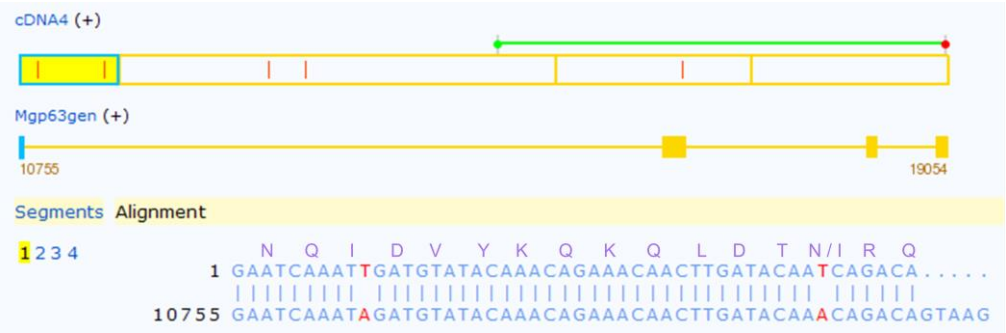
```
  512 TTATAGGGCATGGTTTTAGAAATTAAGAATTTTAGAGGATAATTTTATATTTGTTTCTTGTTTTTGAATG
      ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
12299 TTATAGGGCATGGTTTTAGAAATCAAGAATTTTAGAGGATAATTTTATATTTGTTTCTTGTTTTTGAATG

  582 ACA-TGGATTGTCTTTTAGTGAAAAGCATTTGAATCTGTTTCAATATGAAAGAAATTTAAGACAAGAGTT
      ||| |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||| ||||||
12369 ACAATGGATTGTCTTTTAGTGAAAAGCATTTGAATCTGTTTCTATATGAAAGAAATTTAAGACGAGAGTT

  651 CTGAATCTCATGTTGTTTTGTTTAGATTTTTTTATTTTTATTTAAACTGAAGCAAAGTCTATCGGCTTTT
      |||||||||||||||||| |||||||||||||||||||       ||| |||| ||||||||||||| |||||
12439 CTGAATCTCATGTTGTTATGTTTAGATTTTTTTATTT------AAATTGAATCAAAGTCTATCGTCTTTT

  721 TATCTTTTGCCTAACAATTTGATATTTATTGGAACTAGATGGTTTTACACAAAAAAAA-GATCTTTTGAA
      |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
12503 TATCTTTTGCCCAACAATTTGATATTTATTGGAACTAGATGGTTTTACACAAAAAAAAAGATCTTTTGAA

  790 GAAATTTTTGATATTTTTGAAGAATCTCTTTTCTTCACAGTTATTGAATAGTGCTTACATTTGTAAATAA
      |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
12573 GAAATTTTTGATATTTTTGAAGAATCTCTTTTCTTCATAGTTATTGAATAGTGCTTACATTTGTAAATAA

  860 TTTTCTATATAATATTTTTTTTTATTTCTTTAATGTGATATGTCTTTGAAATGTGCTATTTGTGTAATAT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
12643 TTTTCTATATAATATTTTTTTTTATTTCTTTAATGTGATATGTCTTTGAAATGTGCTATTTGTGTAATAT

  930 ATAATAATTATTGGATATATTTATTTAACTGATGTAAATATTTTAAGGGTTATTTGTGAAGCATTGCAAT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
12713 ATAATAATTATTGGATATATTTATTTAACTGATGTAAATATTTTAAGGGTTATTTGTGAAGCATTGCAAT

 1000 AAGATGTTTCTGTTTATTAACCTATAGTTGGTAGCATTTGTATAATGCCATTCAAGAAAATAATAAAGTA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
12783 AAGATGTTTCTGTTTATTAACCTATAGTTGGTAGCATTTGTATAATGCCATTCAAGAAAATAATAAAGTA

 1070 ATGAAATTAAA
      |||||||||||
12853 ATGAAATTAAA
```

(C)

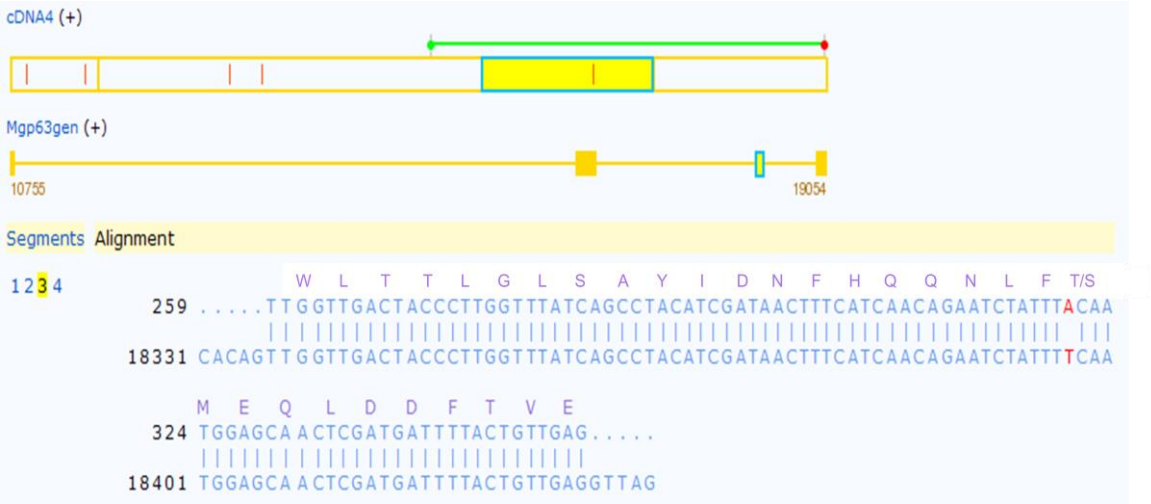

cDNA3 (+)
Mgp63gen (+)
3549
4811
Segments Alignment
1 2 3
M S Q A S V S T T C T
78 .....ATTTGGAAAGTTCACTCATCATCACCCCATAATGTCACAAGCTTCAGTTTCAACTACATGCACAC
4381 TGCAGATTTGGAAAGTTCACTCATCATCACCCCATAATGTCACAAGCTTCAGTTTCAACTACATGCACAC
P S G P P M S Q E T F E Y L W N T L G E V T Q E
143 CATCAGGCCCACCAATGAGCCAGGAAACATTTGAATATTTATGGAATACTCTGGGAGAAGTCACACAAGA
4451 CATCAGGCCCACCAATGAGCCAGGAAACATTTGAATATTTATGGAATACTCTGGGAGAAGTCACACAAGA
G
213 AGG.....
4521 AGGGTAAG

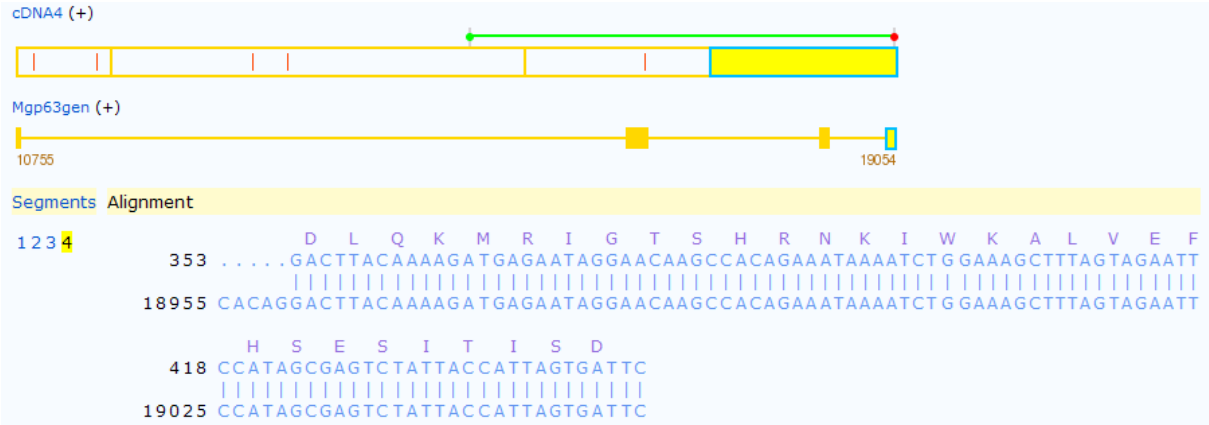

cDNA3 (+)
Mgp63gen (+)
3549
4811
Segments Alignment
1 2 3
G Y T N I T S K E S I D Y A F S E A E D E
216 .....GGGATACACAAACATCACATCTAAAGAATCAATCGACTATGCATTCAGTGAAGCTGAGGATGAAA
4714 TAAAGGGGATACACAAACATCACATCTAAAGAATCAATCGACTATGCATTCAGTGAAGCTGAGGATGAAA
T S I S V E K Y R
281 CCAGTATATCTGTAGAGAAATACAGGAT
4784 CCAGTATATCTGTAGAGAAATACAGGAT

(D)


cDNA4 (+)
Mgp63gen (+)
10755
19054
Segments Alignment
1 2 3 4
N Q I D V Y K Q K Q L D T N/I R Q
1 GAATCAAATTGATGTATACAAACAGAAACAACTTGATACAATCAGACA.....
10755 GAATCAAATAGATGTATACAAACAGAAACAACTTGATACAAACAGACAGTAAG

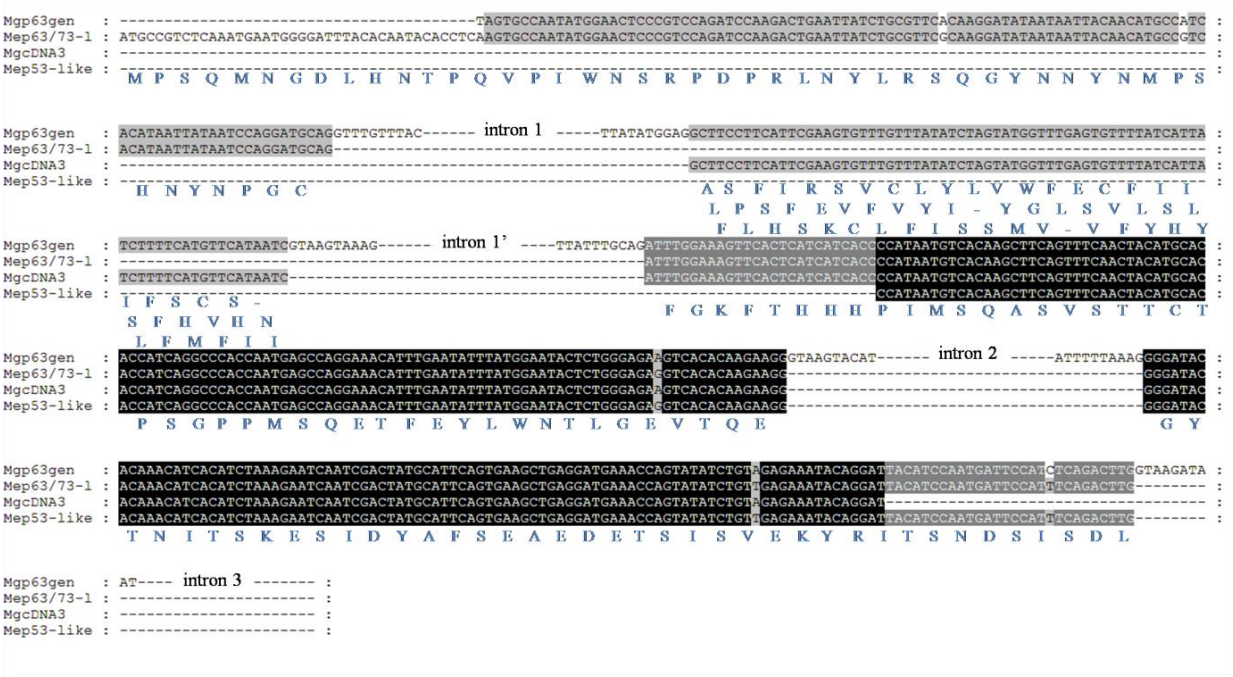
cDNA4 (+)
Mgp63gen (+)
10755
19054
Segments Alignment
1 2 3 4
S L P T T S T A R V V T L P T H D N T P I T
49 .....ATCATTACCAACAACATCAACAGCAAGGGTTGTTACTCTTCCTACTCACGACAACACACCAATAA
16505 TGCAGATCATTACCAACAACATCAACAGCAAGGGTTGTTACTCTTCCTACTCACGACAACACACCAATAA
I H/Q G E G R Q T T L P F T A D L N G Q V T S S
114 CTATACACGGTGAAGGTCGGCAGACGACACTACCATTCACTGCAGACTTAAATGGTCAGGTGACAAGTTC
16575 CTATACAGGGTGAAGGTCGGCAGACTACACTACCATTCACTGCAGACTTAAATGGTCAGGTGACAAGTTC
Q N G V V E N H G N I K E E L M A N G D H S I
184 ACAAAATGGTGTGGTGGAAAATCATGGAAATATTAAAGAAGAACTGATGGCTAATGGGGACCATTCTATT
16645 ACAAAATGGTGTGGTGGAAAATCATGGAAATATTAAAGAAGAACTGATGGCTAATGGGGACCATTCTATT
S N
254 TCAAA.....
16715 TCAAAGTATG


cDNA4 (+)
Mgp63gen (+)
10755
19054
Segments Alignment
1 2 3 4
W L T T L G L S A Y I D N F H Q Q N L F T/S
259 .....TTGGTTGACTACCCTTGGTTTATCAGCCTACATCGATAACTTTCATCAACAGAATCTATTTACAA
18331 CACAGTTGGTTGACTACCCTTGGTTTATCAGCCTACATCGATAACTTTCATCAACAGAATCTATTTTCAA
M E Q L D D F T V E
324 TGGAGCAACTCGATGATTTTACTGTTGAG.....
18401 TGGAGCAACTCGATGATTTTACTGTTGAGGTTAG



Slika 15.A-D Egzoni gena p63 *M. galloprovincialis* utvrđeni sravnjenjem genomskog i kodirajućih sljedova (cDNA1-4, **A-D**) koristeći program Splign dostupan na mreži. Svaki pojedini egzon prikazan je shematski na cDNA i genomskom slijedu kao jedan segment (žuti pravokutnik uokvirenim plavom bojom). Ispod shematskog prikaza, prikazan je nukleotidni slijed egzona pripadajućeg cDNA i genomskog slijeda i proteinski slijed egzona. Nukleotidne razlike između cDNA i genomskog slijeda istaknute su crvenom bojom slova nukleotida.

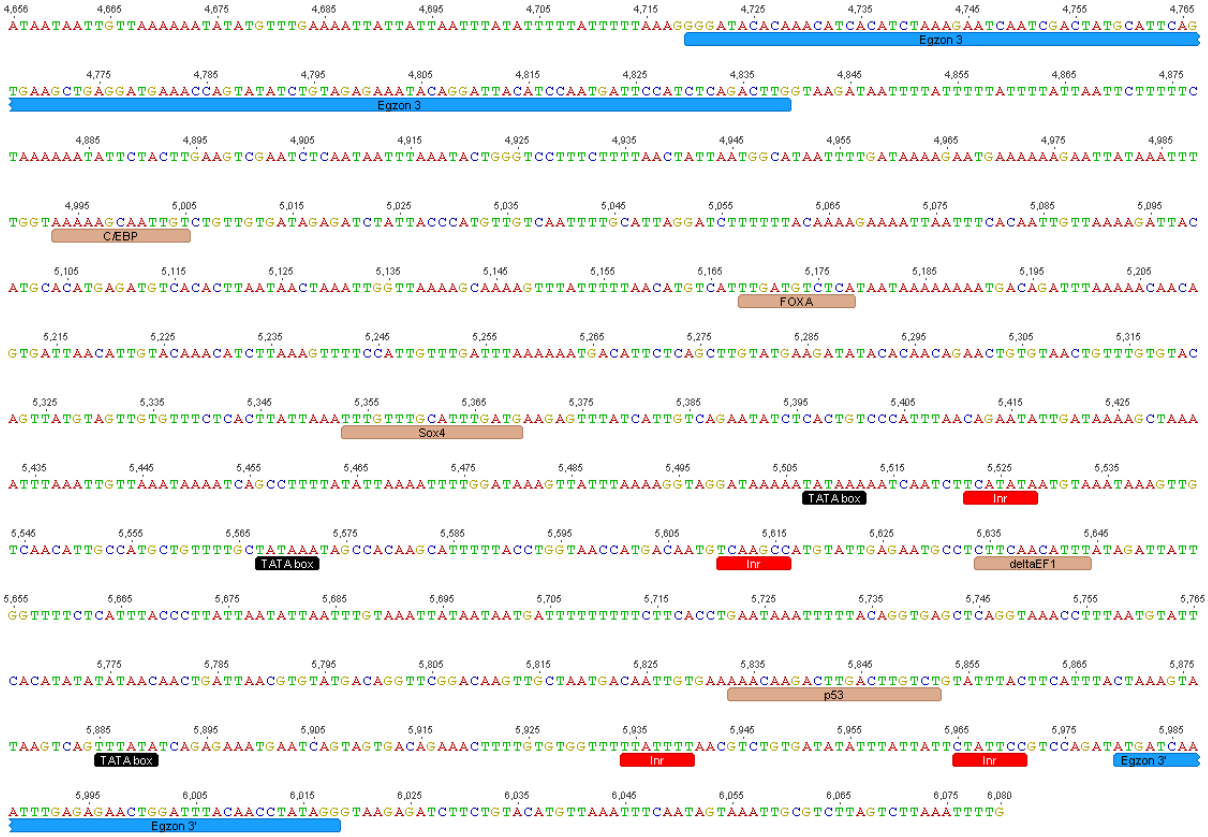
Kodirajući slijed *M. galloprovincialis* cDNA4 određuje posljednjih 40-ak nukleotida egzona 10, cjelovite egzone 11 i 12 te početnih 94 nukleotida egzona 13. Sravnjenje sa „p63/73-like“ *M. edulis* i *M. trossulus* kodirajućim sljedovima otkriva cjeloviti slijed posljednjeg egzona (egzon 13) u strukturi gena porodice p53 *M. galloprovincialis*. Egzon/intron granice gena porodice p53 *M. galloprovincialis* su u skladu sa konsenzus sljedovima izrezivanja introna (GT/AG). Dužine egzona su između 39 i 285 nukleotida, a introna između 195 i 3642 nukleotida (Tablica 5). Uz najduži intron 10' (3642 nt) i intron 1 je vrlo dug (3442 nt) kao što je karakteristično za genske strukture svih članova porodice p53 čovjeka. Na 3' kraju genomski slijed završava sa 1376 nt dugom netranslatiranom regijom (3'UTR). Obje UTR regije gena porodice p53 sadrže poliadenilacijski signal (AATAA) na položaju nt 20491 kod 3'UTR-a i nt 12844 kod unutrašnjeg UTR-a. Dodatno je još kod unutrašnjeg UTR-a prisutan i očuvani motiv K-box (cTGTGATa) na položaju nt 12676 genomskog slijeda. U 3'UTR-u prisutno je više mogućih mjesta vezanja proteina porodice Musashi (ATAGT na 20033-20037, 20067-20071 i 20172-20176 nt te ATTAGT na 20107-20112 i 20372-20377 nt genomskog slijeda). Kao što je već u poglavlju 4.1. utvrđeno, analizom cDNA1 i cDNA2 sljedova, unutrašnji UTR (tj. UTR 10' egzona) ima također očuvana mjesta vezanja proteina porodice Musashi.



Slika 16. Višestruko sravnjenje 5' krajeva genomske DNA *M. galloprovincialis* (Mgp63gen, KP233801) i kodirajućih sljedova *M. galloprovincialis* (MgcDNA3) i *M. edulis* (Mep63/73-like, DQ865150 i Mep53-like, AY579472). Naznačena su mjesta intronskih regija i izvedeni proteinski sljedovi egzona.

Na osnovu vrlo visoke očuvanosti nukleotidnog slijeda, pretraživanjem sa 50-ak nukleotida 5' kraja kodirajućih sljedova *M. edulis* i *M. trossulus* (DQ865151, DQ060436, DQ865153 i DQ060438) na genomskom slijedu porodice p53 *M. galloprovincialis* pronađen je još jedan egzon. Smješten je unutar introna 3 genomskog slijeda i kodira za 13 visoko očuvanih aminokiselina. Ovaj *M. galloprovincialis* egzon odgovara 3' egzonu p63 i p73 gena čovjeka**,** a označava N-kraj ΔNp63 i ΔNp73 proteinskih izoformi pa je time potvrđena mogućnost nastanka ΔN proteinskih izoformi kod dagnje *M. galloprovincialis*. Na osnovu sravnjenja sa *M. edulis* i *M. trossulus* ΔN kodirajućim sljedovima nije moguće odrediti dužinu 5' UTR-a ΔN izoformi *M. galloprovincialis* jer kodirajući sljedovi na 5' kraju započinju startnim kodonom (ATG). Promotorska regija unutar introna 3 sa koje je inicirana transkripcija ΔN izoformi također je nepoznata. Analizom cjelokupne regije introna 3 (4840-5979 nt genomskog slijeda; uzvodno od početnog kodona (ATG) egzona 3' do kraja egzona 3 (GT očuvani početak introna 3) pronađeni su mogući očuvani elementi osnovnog promotora (engl. *core promoter*) eukariotskog gena: TATA box (TATAAA na 5507-5512 nt genomskog

slijeda) i inicijator (TCATATA na 5522-5528 nt) (Slika 17). Potencijalni inicijatorski slijed u odnosu na očuvani konsenzus slijed eukariota (PyPyAN(T/A)PyPy) ima samo jedan različit nukleotid (A na mjestu +4 umjesto pirimidina u konsenzusu eukariota). Inicijator (TCAAGCC na 5610-5616 nt) drugog mogućeg TATA box (TATAAA na 5567-5572 nt) također zadovoljava očuvani konsenzus slijed eukariot; ima samo jedan različit nukleotid (G na mjestu +2 umjesto (T/A) u konsenzusu eukariota). Treći mogući TATA box (TTTATA na 5885-5890 nt) ima dva moguća inicijatora (TTATTTT na 5934-5940 nt i CTATTCC na 5965-5971 nt) koji oba u potpunosti zadovoljavaju očuvani konsenzus slijed eukariota. TATA box i inicijator (točnije, A na trećoj poziciji inicijatora kao mjesto početka transkripcije) su u svim opisanim slučajevima na očekivanoj međusobnoj udaljenosti (~20-80 nt). U analiziranoj regiji introna 3 pronađen je i potencijalni CAAT box (4993-5005 nt genomskog slijeda) te očuvani motivi vezanja transkripcijskih faktora: FOXA (5168-5178 nt), Sox4 (5353-5369 nt), deltaEF1(5634-5644 nt) i p53 (5833-5852) (Slika 17). Očuvano vezno mjesto p53 (RE) ima strukturu dva uzastopna ponavljajuća motiva od 10 pb bez razmaka među njima. U odnosu na općeprihvaćeni konsenzus ponavljajućih motiva („RRRCWWGYYY“), pronađeno p53-RE mjesto razlikuje se u dva nukleotida; A (nt 5840) umjesto Y u prvom te T(nt 5843) umjesto R u drugom ponavljajućem motivu.

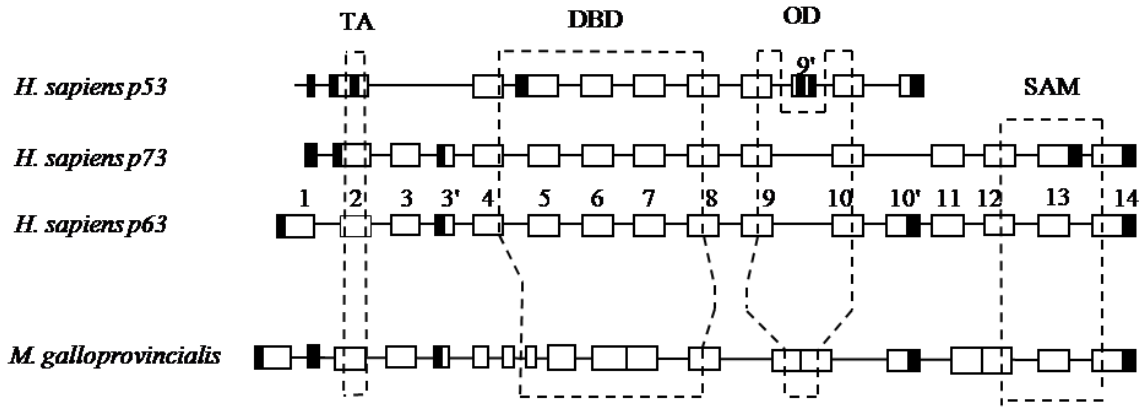


Slika 17. Prikaz regije genomskog slijeda (4656-6080 nt) porodice p53 *M. galloprovincialis* sa označenim egzonima (3 i 3'), elementima osnovnog promotora (TATA box i inicijator (Inr)) i očuvanim mjestima vezanja transkripcijskih faktora (deltaEF1, Sox4, FOXA ,C/EBP i p53).

Ranije poznati, djelimični genomski slijed porodice p53 *M. galloprovincialis* (AJ66664), dužine 1333 nt, gotovo je identičan regiji 8493- 9825 nt (početak introna 6 do sredine egzona 8) dobivenog cjelovitog genomskog slijeda *M. galloprovincialis* (KP233801). Uočene razlike (polimorfizmi) između ova dva genomska slijeda su promjene jednog nukleotida (SNP) na tri mjesta (genomski slijed; nt 9027 A/T, nt 9130 A/T i nt 9464 T/G) i gubitak (delecija) nukleotida (nt 9409 G), sve u intronskim regijama gena.

Ukupna egzon-intron organizacija prisutnog gena porodice p53 kod dagnje *M. galloprovincialis* sličnija je p63 i p73 homolozima, nego p53 homologu čovjeka. Genomske strukture gena p63 i p73 čovjeka vrlo su međusobno slične. Oba gena kodiraju dva 3'UTR-a, od kojih je udaljeniji smješten na samom kraju gena. Jasna strukturalna razlika između p63 i p73 odnosi se na smještaj drugog, bližeg 3'UTR-a, koji je kod p63 smješten na kraju egzona 10', a kod p73 na kraju egzona 13 . Na taj način, regija između egzona 10 (posljednji egzon

očuvane središnje regije prisutne kod svih proteina porodice p53) i egzona sa tim bližim 3'UTR-om sadrži kod p63 samo jedan intron (intron 10), a kod p73 tri introna (introni 10, 11 i 12) i dva egzona (egzoni 11 i 12). Egzon-intron struktura ove genomske regije može zato poslužiti za razlikovanje p63 od p73. Dobiveni cjeloviti genomski slijed *M. galloprovincialis* nema dodatnih egzona između očuvanog egzona 10 i egzona sa 3'UTR-om pa prema strukturi 3' kraja gen porodice p53 *M. galloprovincialis* odgovara genu p63 čovjeka (Slika 18). U usporedbi sa p63 čovjeka (Gene ID 8626), *M. galloprovincialis* p63 genu nedostaju tri introna, kod čovjeka naznačeni kao introni 6, 9 i 11. U regiji naznačenoj kao egzon 4 čovjeka, *M. galloprovincialis* p63 ima dva dodatna introna te na taj način egzoni 4, 5 i 6 dagnje zajedno odgovaraju egzonu 4 čovjeka. Svi ostali položaji i faze introna su istovjetne kod struktura p63 gena dagnje i čovjeka (Slika 18).



Slika 18. Shema sravnjenja gena porodice p53 čovjeka (p53, p63 i p73) i gena porodice p53 *M.galloprovincialis*. Egzoni su prikazani pravokutnicima (crno su označeni UTR-ovi), a introni ravnim linijama. Očuvane domene porodice p53 (TA, DBD, OD i SAM) uokvirene su iscrtkanim linijama.

Tablica 5 Egzon-intron organizacija gena porodice p53 *M.galloprovincialis*. Dužina egzona 1 je od prvog nukleotida genomskog slijeda *M. galloprovincialis* do egzon-intron granice. Dužina egzona 3' je od startnog kodona (ATG) do egzon-intron granice. Dužine egzona 10' i 13 su od egzon-intron granica do stop kodona (TGA). Izdvojeno su naznačene dužine netranslatiranih regija (UTR) egzona.

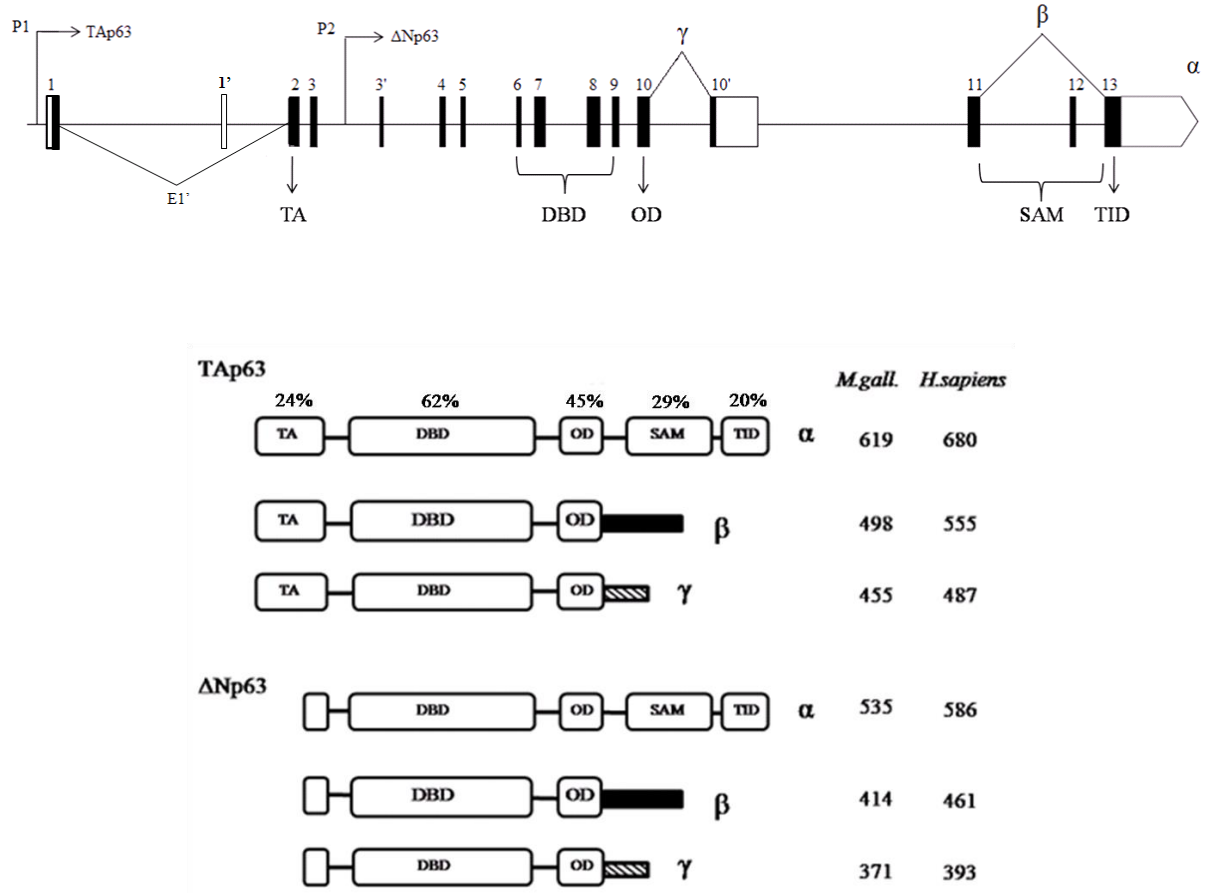
Egzon/intron	Dužina egzona (nt)	Dužina introna (nt)	Faza introna
1	106	3442	2
1'	77	760	2
2	138	195	2
3	121	1140	0
3'	39	1028	0
4	99	275	0
5	71	923	2
6	79	243	0
7	187	765	1
8	217	235	2
9	116	335	1
10	211	1081	2
10'	97+887 UTR	3642	2
11	210	1616	2
12	94	530	0
13	285 + 1376 UTR	/	/

4.4. Očekivane ekspresijske izoforme gena porodice p53 *M. galloprovincialis*

Prema utvrđenoj strukturi gena porodice p53 *M.galloprovincialis* određene su očekivane ekspresijske izoforme gena. Strukture i veličine svih pretpostavljenih proteinskih izoformi dagnje odgovaraju izoformama gena p63 čovjeka (Slika 19). Početkom transkripcije od egzona 3' (5980-6018 nt genomskog slijeda *M.galloprovincialis*) nastaju sljedovi mRNA koji kodiraju ΔN proteinske izoforme N-kraja gena porodice p53 *M. galloprovincialis*. Karakteristični ΔN fragment N-kraja proteina kodiran egzonom 3' dužine je 13 aminokiselina. Očuvana struktura gena podupire postojanje i tri izoforme C-kraja (α, β i γ). mRNA molekula ΔNp63α izoforme dužine je 2984 nt s otvorenim okvirom čitanja od 535 ak (od ATG start kodona na 1-3 nt do TGA stop kodona na 1606-1608 nt mRNA). mRNA molekula ΔNp63β izoforme dužine je 2890 nt s otvorenim okvirom čitanja od 414 ak (od ATG start kodona na 1-3 nt do TGA stop kodona na 1243-1245 nt mRNA). mRNA molekula ΔNp63γ izoforme

dužine je 2003 nt s otvorenim okvirom čitanja od 371 ak (od ATG start kodona na 1-3 nt do TGA stop kodona na 1114-1116 nt mRNA). Dok kodirajući sljedovi ΔN proteinskih izoformi imaju samo jedan mogući početak transkripcije, TA izoforme imaju tri moguća početaka transkripcije; od egzona 1, egzona 1' i egzona 2. Početkom transkripcije od egzona 1, uz alternativno izrezivanje introna 5' kraja gena (gubitak egzona 1'), nastaje mRNA molekula TAp63α izoforme dužine 3310 nt s otvorenim okvirom čitanja od 619 ak (od ATG start kodona na 75-77 nt do TGA stop kodona na 1932-1934 nt mRNA). mRNA molekula TAp63β izoforme dužine je 3216 nt s otvorenim okvirom čitanja od 498 ak (od ATG start kodona na 75-77 nt do TGA stop kodona na 1569-1571 nt mRNA). mRNA molekula TAp63γ izoforme je dužine 2329 nt s otvorenim okvirom čitanja od 455 ak (od ATG start kodona na 75-77 nt do TGA stop kodona na 1440-1442 nt mRNA). Početkom transkripcije od egzona 2 dužine nastalih molekula mRNA (TAp63α, TAp63β i TAp63γ) kraće su za 106 nt 5' kraja (egzon 1) u odnosu na navedene dužine istih mRNA izoformi nastalih transkripcijom od egzona 1. Dva su moguća početka (ATG start kodona) otvorenog okvira čitanja sa odgovarajućim konsenzus Kozak slijedom ovih mRNA molekula: jači na 32-34 nt i slabiji na 80-82 nt. U oba slučaja na +1 poziciji Kozak slijeda nije G, a kod jačeg startnog kodona na -3 poziciji Kozak slijeda je A. Koristeći jači ATG kao početak otvorenog okvira čitanja, veličine nastalih proteinskih izoformi su: TAp63α 598 ak, TAp63β 477 ak i TAp63γ 434 ak. Početkom transkripcije od egzona 1' nastale mRNA proteinskih izoformi TAp63α, TAp63β i TAp63γ duge su istim redosljedom 3281, 3187 i 2300 nt. Otvoreni okviri čitanja ovih transkripata od egzona 1' jednaki su onima transkripata od egzona 2 jer je egzon 1' nekodirajući. Izvedeni cjeloviti proteinski sljedovi svih 6 očekivanih izoformi gena porodice p53 *M. galloprovincialis* (ORF od egzona 1) navedeni su u Prilogu 1.

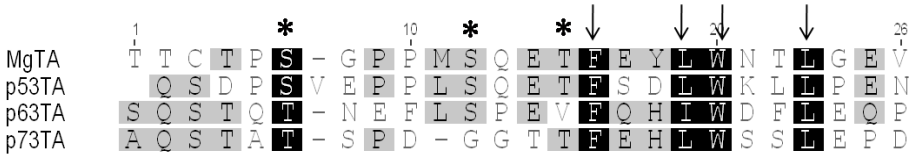
δ i ε izoforme C-kraja proteina p63 čovjeka ne mogu se sa sigurnošću predvidjeti kod dagnje *M. galloprovincialis*. p63δ nastaje alternativnim izrezivanjem egzona 12 i 13, a egzon 12 čovjeka odgovara jednom dijelu egzona 11 *M. galloprovincialis* pa nije sigurno kako p63δ može nastati kod dagnje *M.galloprovincialis*. p63ε izoformu također nije moguće sa sigurnošću predvidjeti kod dagnje *M. galloprovincialis* jer ova izoforma nastaje prijevremenom terminacijom transkripcije u intronu 10, a mnogo je mogućih uzroka za takav događaj.

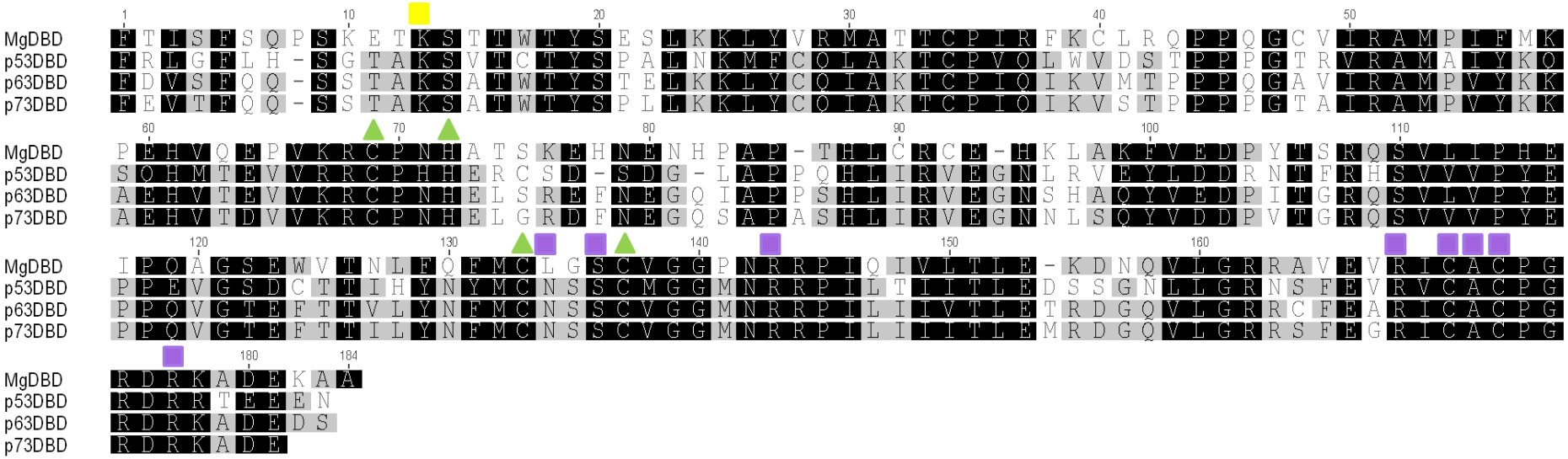


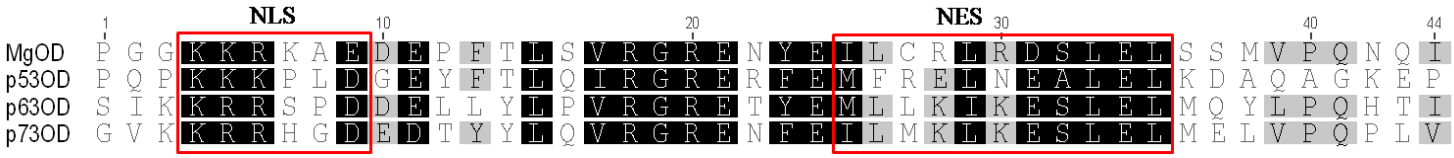
Slika 19. Shematski prikaz strukture gena i očekivanih proteinskih izoformi porodice *p53 Mytilus galloprovincialis*. **(A)** Egzon-intron struktura gena; numerirani pravokutnici označavaju egzone (bijelo su netranslatirane regije), a linije koje ih povezuju introne. Naznačena su alternativna promotorska mjesta (P1 i P2) i produkti alternativnog izrezivanja (α, β i γ). Očuvane domene su označene; transaktivacijska (TA), DNA-veznaća (DBD), oligomerizacijska (OD), sterilnog α motiva (SAM) i transinhibitorna. **(B)** Očekivane proteinske izoforme. Navedene su veličine (broj aminokiselina) proteinskih izoformi i uspoređene sa veličinama odgovarajućih proteinskih izoformi proteina p63 čovjeka. Iznad najduže proteinske izoforme (TAp63α) navedene su identičnosti (%) očuvanih funkcionalnih domena p63 čovjeka (GeneID 8626) i dagnje *M.galloprovincialis*.

Očuvane funkcionalne domene porodice p53 prisutne su kod pretpostavljenih sljedova proteinskih izoformi dagnje *M. galloprovincialis* (Slika 19 i 20). Transaktivacijska domena (TA) je kodirana sa egzona 2 dagnje i sa TA domenama p53(NP_000537), p63 (NP_003713) i p73 (NP_005418) čovjeka dijeli redom 44%, 24% i 32% identičnosti. TA domena proteina *M. galloprovincialis* sa članovima porodice p53 čovjeka dijeli očuvana mjesta ključna za interakciju sa Mdm2 proteinom. U proteinskom slijedu TAp63α izoforme *M. galloprovincialis* (619 ak) to su aminokiseline: F43, L46, W47 i L50 (Slika 20). TA domena

proteina *M. galloprovincialis* ima očuvana mjesta fosforilacije TA domene p53 čovjeka. U proteinskom slijedu TAp63α izoforme *M. galloprovincialis* (619 ak) to su aminokiseline: S34, S39 i T42 (Slika 20). Na C-kraju pretpostavljenih proteinskih sljedova dagnje *M. galloprovincialis* prisutne su oligomerizacijska (OD) i SAM domena. OD je kod *M. galloprovincialis* kodirana sa egzona 10 i sa OD proteina p53, p63 i p73 čovjeka dijeli redom 36%, 45% i 45% identičnosti. Unutar OD je kod *M. galloprovincialis* očuvan i slijed (392-402 ak TAp63α izoforme) signala za izlaz iz jezgre (engl. *Nuclear export signal*, NES) p53 proteina čovjeka. Na samom početku OD *M. galloprovincialis* je i očuvani slijed (371-376 ak TAp63α izoforme) signala za smještanje u jezgru (eng. *Nuclear localisation signal*, NLS). SAM je kodirana egzonima 12 te dijelom 11 i 13. Sa SAM domenama proteina p63 i p73 čovjeka dijeli redom 29% i 32% identičnosti. Najveći stupanj identičnosti sa tri paraloga porodice p53 čovjeka izoforme *M.galloprovincialis* dijele u DNA-veznoj domeni; 41% sa p53, 62% sa p63, a 58% sa p73. Ta centralna domena je kod *M. galloprovincialis* kodirana egzonima 7 i 8 te dijelom 6 i 9. Četiri aminokiseline DBD porodice p53 koje ostvaruju interakciju sa ionom cinka su očuvane i kod DBD *M. galloprovincialis* (C231, H234, C293 i C297 izoforme TAp63α *M. galloprovincialis*). Osam aminokiselina DBD porodice p53 važnih za vezanje na DNA su očuvane i kod DBD *M. galloprovincialis* (L294, S296, R303, R327, C329, A330, C331 i R336 izoforme TAp63α *M. galloprovincialis*). Kod DBD *M. galloprovincialis* je očuvan i lizin (K175), mjesto acetilacije p53 čovjeka. Pretraga očuvanih domena alatom NCBI Batch CD-Search ne prepoznaje transinhibitornu domenu (TI) na izvedenom proteinskom slijedu gena porodice p53 *M. galloprovincialis*. Sravnjenje C-kraja (57 ak) α izoformi porodice p53 *M. galloprovincialis* sa sljedovima TI domena p63 i p73 proteina čovjeka (zadnjih 70 ak C-kraja α izoformi) upućuju na prisutnost TI domene i kod *M. galloprovincialis* jer je utvrđena identičnost ovih regija od oko 20% (Slika 20). TID *M. galloprovincialis* kodirana je sa egzona 13. C-kraj izoformi *M. galloprovincialis* ne pokazuje nikakvu sličnost sa baznim područjem p53 proteina čovjeka.


MgTA
p53TA
p63TA
p73TA


MgDBD
p53DBD
p63DBD
p73DBD


NLS
NES
MgOD
p53OD
p63OD
p73OD

```
           1                   10                  20                  30                  40                  50
MgSAM      A N G D H S I S N W L T T L G L S A Y I D N F H Q Q N L F S M E Q L D D F T V E D L Q K M R I G T S H R N K I W K A L
p63SAM     Y P T D C S I V S F L A R L G C S S C L D Y F T T Q G L T T I Y Q I E H Y S M D D L A S L K I P E Q F R H A I W K G I
p73SAM     Y H A D P S L V S F L T G L G C P N C I E Y F T S Q G L Q S I Y H L Q N L T I E D L G A L K I P E Q Y R M T I W R G L
           60          65
MgSAM      V E F H S E
p63SAM     L D H R Q L
p73SAM     Q D L K Q G

           1                   10                  20                  30                  40                  50
MgTID      D V S T A S T I S M T S Q A S I S Q N S Q N S T Y C P G Y Y E V T R Y T F K H T V S Y T K V - - - - - - - D E R S P
p63TID     F S S P S H L L R T P S S A S T V S V G S S E T R G E R V I D A V R F T L R Q T I S F P P R - - - - - - - D E W N D
p73TID                               A A T I S I G G S G E L Q R Q R V M E A V H F R V R H T I T I P N R G G P G G G P D E W A D
           60                  70                  80      83
MgTID      K R A K V D
p63TID     F N F D M D A R R N K Q Q R I K E E G E
p73TID     F G F D L P D C K A R K Q P I K E E F T E A E I H
```

Slika 20. Sravnjenje proteinskih sljedova očuvanih funkcionalnih domena porodice p53 kod dagnje *M. galloprovincialis* i tri paraloga čovjeka (p53, p63, p73). Zvjezdice (*) označavaju mjesta fosforilacije TA domene, a strjelicama su označene četiri aminokiseline TA domene neophodne za vezanje Mdm2 proteina. Zelenim trokutima su označene aminokiseline DBD koje ostvaruju interakciju sa ionom cinka, ljubičasti kvadrati označavaju mjesta važna za vezanje na DNA, a žuti kvadrat mjesto acetilacije. Na OD su crveno uokvireno naznačeni aminokiselinski sljedovi signala za smještanje u jezgru (NLS) i eksport iz jezgre (NES).

4.5. Intronske regije

Sa ciljem pronalaska mogućih homologija sa sljedovima ponavljajućih regija, pokretnih genetičkih elemenata, okvira čitanja (regulatornih) proteina ili tRNA, svi intronski sljedovi p63 gena *M. galloprovincialis* analizirani su koristeći BLAST pretraživanje. Samo je BLASTx pretragom NCBI baze podataka utvrđena zanimljiva sličnost visoke homologije (75% identičnost) pretpostavljenog proteinskog slijeda regije introna 1 (2964-2893 nt u odnosu na početak genomskog slijeda) gena p63 *M. galloprovincialis* i početne 24 aminokiseline Wrap73 proteina čovjeka (NP_060288) (Slika 21). Ovih početnih 24 aminokiseline je cjeloviti izvedeni proteinski slijed egzona 1 gena Wrap73 čovjeka (Gene ID 49856), utvrđeno prema opisanoj egzon-intron strukturi gena Wrap73 čovjeka (13 egzona).

WD repeat-containing protein WRAP73 [Homo sapiens]
Sequence ID: ref|NP_060288.3| Length: 460 Number of Matches: 1
▶ See 5 more title(s)

Range 1: 1 to 24 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
42.7 bits(99)	2.5	Compositional matrix adjust.	18/24(75%)	21/24(87%)	0/24(0%)	-3

```
Query  2858  MNFSELFRQTSQLCKFSPDGKYLV  2787
             MNFSE+F+ +S LCKFSPDGKYL
Sbjct  1     MNFSEVFKLSSLLCKFSPDGKYLA  24
```

Slika 21. Prikaz rezultata BLASTx pretrage GenBank baze podataka koristeći slijed introna 1 gena p63 *M.galloprovincialis* kao upit u pretraživanju.

Wrap73 je protein koji sadrži WD ponavljanja (strukturalni motiv od oko 40 aminokiselina, triptofan (W) i aspartat (D) kao krajnje aminokiseline), a kod čovjeka je ovaj protein kodiran genom smještenim uzvodno od p73 gena (GeneID 7161) na kromosomu 1, ali kao „antisense“ p73. Pretpostavljeni proteinski slijed regije introna 1 gena p63 *M. galloprovincialis* je kodiran kao „antisense“ gena p63 *M. galloprovincialis*.

Na cjelovitom genomskom slijedu p63 gena *M. galloprovincialis* pronađeno je također, koristeći program Tandem Repeats Finder, ukupno 6 regija uzastopnih ponavljanja. Svih 6 regija je u područjima introna. Najveći broj regija ponavljanja (3) je u intronu 10', na pozicijama 13045-13084 nt, 14503-14620 nt i 14742-15234 nt u odnosu na početak genomskog slijeda. Jedna regija ponavljanja, na poziciji 10339-10405 nt u odnosu na početak

genomskog slijeda, smještena je u intronu 9, a jedna u intronu 11, na poziciji 17282 – 18180 nt. Veličina ponavljanog slijeda (perioda) i broj ponavljanja (broj kopija) različiti su za svaku od 6 regija. Veličine perioda kreću se u rasponu od 19 do 65 nukleotida te tako pripadaju minisatelitima; dužim uzastopno ponavljajućim sljedovima. Rezultati analize uzastopnih ponavljanja su opisani u Tablici 6. Udio identičnih nukleotida svih kopija određenog perioda (identičnost) kreće se u rasponu od 88-100 %. U Tablici 6 naznačeni su i udjeli svakog od nukleotida (A,T,G,C) u ponavljanom slijedu (periodu).

Tablica 6 Rezultat pretraživanja uzastopnih ponavljanja genomskog slijeda koristeći program Tandem repeats finder.

Pozicija (nt)	Intron	Veličina perioda (nt)	Broj kopija	Identičnost (%)	A (%)	C (%)	G (%)	T (%)
10339-10405	9	33	2.0	94	44	8	13	32
13045-13084	10'	19	2.1	100	22	30	10	37
14503-14620	10'	53	2.2	93	36	17	22	23
14742-15234	10'	51	9.7	88	38	24	15	21
17282-18180	11	65	13.8	99	40	15	13	30

4.6. Usporedba strukture gena porodice p53 mekušaca

Sravnjenje jedinog cjelovitog genomskog slijeda porodice p53 mekušaca, slijeda školjkaša *Mya arenaria* („p63/p73 and p53 gene“, FJ041332) i cjelovitih kodirajućih sljedova porodice p53 istog školjkaša (AF253323 i AF253324), otkrilo je strukturu gena porodice p53 *Mya arenaria*. Kao i kod *M. galloprovincialis*, struktura gena *M. arenaria* odgovara strukturi gena p63 čovjeka (Slika 22). Jedini egzon koji nije otkriven ovim sravnjenjem je egzon koji odgovara egzonu 3' gena p63 i p73, prvi egzon ΔN proteinskih izoformi. BLASTn pretraga GenBank baze, koristeći *M. arenaria* genomsku regiju između egzona koji odgovaraju egzonima 3 i 4 čovjeka (nt. 1722-6046 slijeda FJ041332), pronalazi vrlo visoku sličnost sa *M. edulis* i *M. trossulus* ΔN frgmentima (prvih 39 nukleotida sljedova DQ865151, DQ865153, DQ060436 i DQ060438) što upućuje na postojanje 3' egzona i kod *M. arenaria*. Pronađeni egzon 3' *M. arenaria* dužine je 39 nt i izvedenog proteinskog slijeda od 13 ak. Glavna razlika

u strukturi gena *M. arenaria* u odnosu na p63 čovjeka vidljiva je u središnjoj regiji gena (homolozi egzona 4-9 čovjeka) gdje nedostaju tri introna dok su dva dodatna introna prisutna unutar regije homologne egzonu 4 čovjeka. DNA-veznaća domena proteina porodice p53 *M. arenaria* kodirana je tako krajem egzona 6, egzonom 7 te početkom egzona 8. Kraj egzona 8 i početak egzona 9 kodiraju oligomerizacijsku domenu (Slika 22). 3' kraj gena p63 *M. arenaria* identičan je genu p63 čovjeka te je uz najdužu α izoformu, mehanizmom alternativnog izrezivanja introna, moguć nastanak i β, γ i δ izoformi 3' kraja. ε izoformu, kao ni kod *M. galloprovincialis*, nije moguće sa sigurnošću predvidjeti kod *M. arenaria*. Većina SAM domene kodirana je egzonom 13 pa izoforme β, γ i δ nemaju ovu domenu jer je ovaj egzon alternativnim izrezivanjem introna izgubljen iz njihovih zrelih mRNA. Prvi egzon gena porodice p53 *M. arenaria* nije kodirajući, nego otvoreni okvir čitanja TA izoformi započinje od ATG start kodona u egzonu 2 (931-933 nt genomskog slijeda *M. arenaria*, FJ041332).

NCBI baza podataka sadrži sastavljene cjelokupne genome četiri predstavnika koljena Mollusca: puževa (Razred Gastropoda) kalifornijski morski zekan *Aplysia californica*, slatkovodni puž *Biomphalaria glabrata*, priljepak *Lottia gigantea* te školjkaša (Razred Bivalvia) pacifička kamenica *Crassostrea gigas*. tBLASTn pretraga sekvenciranih genoma aminokiselinskim slijedom DNA-vezne domene proteina porodice p53 *M. galloprovincialis* pronašla je samo po jedan p53 sličan slijed po haploidnom genomu svakog od četiri predstavnika koljena Mollusca. Pronađeni sljedovi porodice p53 korišteni su za analizu strukture gena porodice p53 u koljenu Mollusca (Tablica 7).

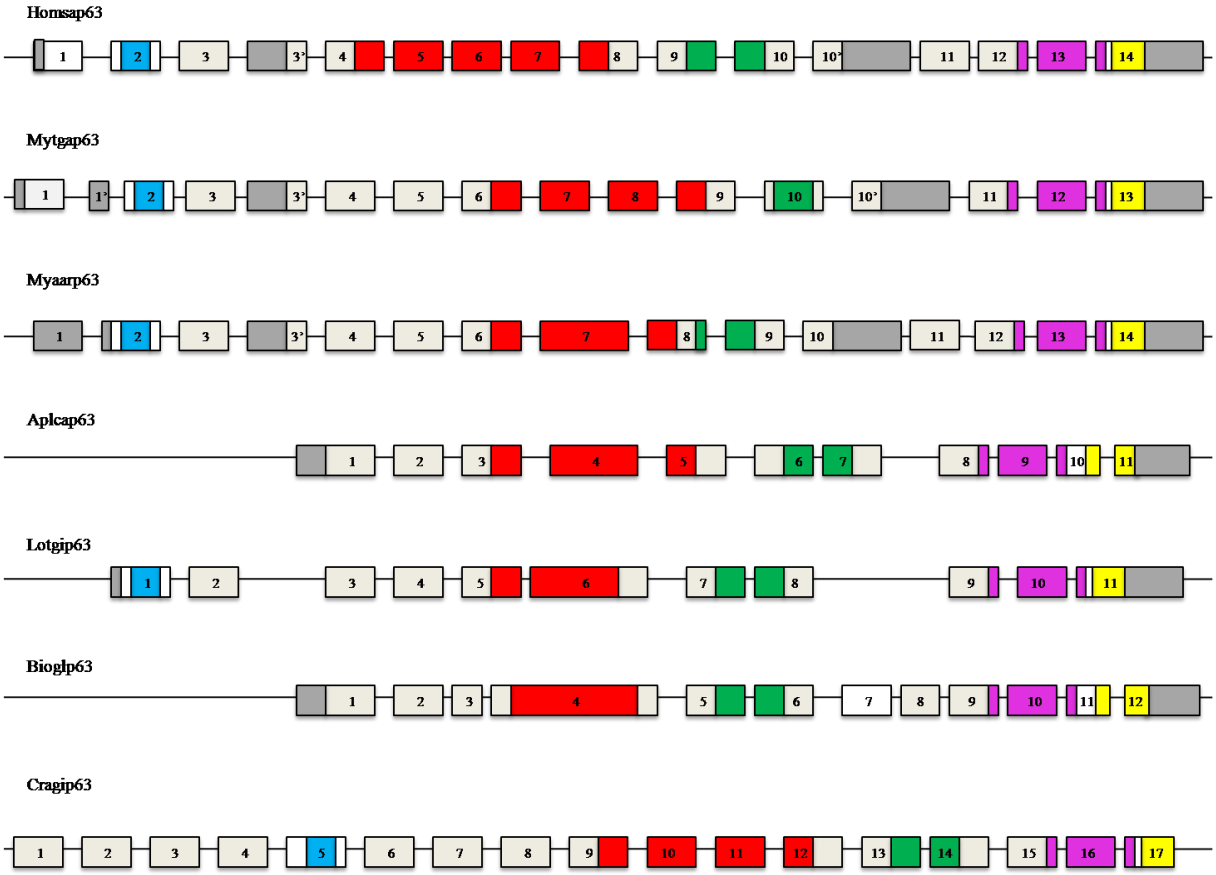
Tablica 7 Porijeklo i glavne karakteristike genomskih sljedova porodice p53 mekušaca dostupnih u GenBank bazi podataka. Navedeni sljedovi korišteni su u usporedbama strukture gena (Slika 22) i proteinskim sravnjenjima (Slika 24).

Naziv vrste	NCBI broj Bioprojekta	Publikacija	NCBI pristupni broj genomskog slijeda	Naziv gena/proteina	Veličina gena (kb)	Veličina proteina (ak)	NCBI pristupni broj proteina
Aplysia californica (Kalifornijski morski zekan)	PRJNA209509	Broad Institute (nije publicirano)	NW_004798815	tumor protein 63-like	18.772	542	XP_005113064
Biomphalaria glabrata (slatkovodni puž)	PRJNA12879	Washington University (nije publicirano)	APKA01029991	Gene BGLTMP012777*	11.214	581	/
Crassostrea gigas (Pacifička (japanska) kamenica)	PRJNA70283	Zhang i sur. (2012)	JH815916	Tumor protein 63	42.200	790	EKC31656
Lottia gigantea (Pacifički priljepak)	PRJNA175706	Simakov i sur. (2013)	AMQO01000525	hypothetical protein LOTGIDRAFT_182533	16.549	607	ESO91245
Mya arenaria (Pješčano zijalo)	/	nije publicirano	FJ041332	p63/p73 and p53 gene	16.798	621**	ACK28180
Mytilus galloprovincialis (Mediteranska dagnja)	/	Baričević i sur. (2015)	KP233801	p63	20.620	619**	AKE31911
Homo sapiens (Čovjek)	PRJNA168	Muzny et al. (2006)	NC_000003 REGION: 189631160…189897279, GPC_000001295	TP63	266.120	680**	NP_003713

*Naziv gena prema podacima na web stranici genomskog projekta *Biomphalaria glabrata* (https://www.vectorbase.org/organisms/biomphalaria-glabrata)

**TAp63α

Za *A. californica* i *C. gigas* dostupne su anotacije gena uz NCBI zapis dok su za *L. gigantea* i *B. glabrata* anotacije gena preuzete iz podataka dostupnih na web stranicama pripadajućih genomskih projekata. *A. californica* i *L. gigantea* geni porodice p53 organizirani su u 11 egzona isprekidani sa 10 introna. *B. glabrata* gen porodice p53 broji 12 egzona i 11 introna. Gen *C. gigas* je znatno duži od ostalih predstavnika koljena Mollusca te broji 17 egzona i 16 introna. Izvedeni proteinski sljedovi gena porodice p53 svih predstavnika Mollusca imaju DNA-veznu, oligomerizacijsku, domenu sterilnog α motiva i transinhibitornu domenu dok je transaktivacijska domena prisutna samo kod *L. gigantea* i *C. gigas*. Školjkaši *M. galloprovincialis* i *C. gigas* dijele gotovo identičnu strukturu gena sa jedinom razlikom u regiji koja kodira za oligomerizacijsku domenu gdje *C. gigs* ima dodatni intron koji dijeli oligomerizacijsku domenu u dva egzona. Genska organizacija oligomerizacijske domene kao kod *C. gigas* prisutna je i kod p63 sljedova puževa *A. californica*, *L. gigantea* i *B. glabrata*. Puževi *L. gigantea* i *B. glabrata* dijele i istu organizaciju regije gena koja kodira za DNA-veznuću domenu, ali razlikuju se u strukturi 3' kraja gena gdje je dodatni intron prisutan kod *B. glabrata* te dijeli transinhibitornu domenu u dva egzona; 11. i 12. egzon gena *B. glabrata* (Slika 22). Kao i kod *M. galloprovincialis* i *M. arenaria*, struktura gena porodice p53 *A. californica*, *B. glabrata*, *C. gigas* i *L. gigantea* odgovara strukturi gena p63 čovjeka (Slike 22 i 23).

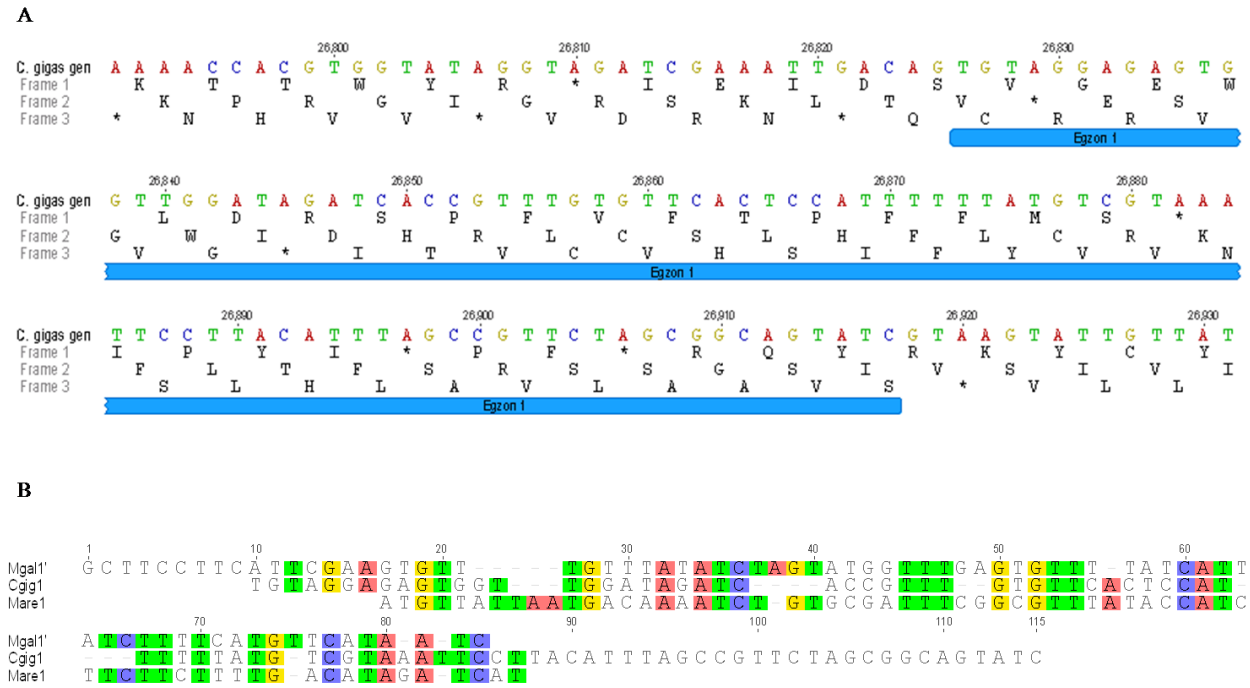


Slika 22. Shematski prikaz egzon-intron struktura p63 gena čovjeka (Homsap63) i mekušaca (Mytgap63- *Mytilus galloprovincialis*, Aplcap63- *Aplysia californica*, Lotgip63- *Lottia gigantea*, Bioglp63- *Biomphalaria glabrata*, Cragip63-*Crassostrea gigas*). Genomske strukture prikazane kako su anotirane na pripadajućim sljedovima sekvenciranih genoma. Pravokutnici prikazuju egzone, ravne linije između njih introne. Očuvane domene porodice p53 označene su različitim bojama (TA-plavo, DBD-crveno, OD-zeleno, SAM-ljubičasto i TID-žuto), a netranslatirane regije (UTR) sivo.

Usporedbom sa izvedenim proteinskim sljedovima ostalih predstavnika Mollusca, protein *C. gigas* ima produženi N-kraj proteina (Slika 24). Ovaj produžetak dug je 191 aminokiselinu i kodiran je sa prva četiri egzona gena *C. gigas* (JH815916). BLASTp pretragom GenBank baze podataka nije utvrđena sličnost ovog aminokiselinskog produžetka s proteinskim sljedovima porodice p53 kao ni značajna sličnost s bilo kojim drugim proteinskim slijedom u GenBank bazi podataka. BLASTx pretragom GenBank baze podataka utvrđena je, kao i za gen porodice p53 *M. galloprovincialis*, sličnost visoke homologije (75% identičnosti) regije introna 4 *C. gigas* sa Wrap73 proteinom čovjeka (NP_060288). U intronu 4 *C. gigas*

pronađeni su tako kodirajući sljedovi egzona 1 i egzona 2 Wrap73 homologa. Koristeći u daljnjim BLASTx analizama regiju genomskog slijeda p63 *C. gigas* uzvodno od introna 4, utvrđen je i preostali dio Wrap73 homologa (egzoni 3,4,5 i 6). Na isti način Wrap73 protein (izvedeni proteinski slijed egzona 1,2,3,4,5 i 7) detektiran je i kod *L. gigantea* u regiji genomskog slijeda uzvodno od prvog anotiranog egzona.

Sravnjenjem proteinskih sljedova, utvrđena je vrlo visoka homologija aminokiselinskih sljedova *C. gigas* egzona 5 i *M. galloprovincialis* egzona 2 te dalje nizvodno. Ove analize su potvrdile da je 5' kraj gena *C. gigas* pogrešno anotiran tako da su zapravo prva četiri egzona netranslatirana (dio 5' UTR-a) ili uopće ne pripadaju ispravnom okviru čitanja gena p63 *C. gigas*. Pronađena je 100% identičnost 5'-kraja kodirajuće DNA *C. gigas* (AM236465) sa egzonom 5 i središnjom regijom egzona 4 gena porodice p53 *C. gigas*. Na ovaj način utvrđen je ispravan prvi egzon gena *C. gigas* (ili barem njegov 3'kraj), smješten u regiji anotiranoj kao intron 4 gena *C. gigas*. U skladu s tim, ispravan otvoreni okvir čitanja počinje unutar egzona 5 gena porodice p53 *C. gigas*. Ispravan prvi egzon *C. gigas* dužine je 92 nt i nije kodirajući (Slika 23). Očuvani konsenzusi izrezivanja introna (GT/AG) omeđuju egzon 1 *C.gigas* (AG ispred, a GT iza egzona) što upućuje da je ovaj ispravan egzon 1 ipak cjeloviti slijed te da ga se kao i kod *M. galloprovincialis* može zapravo nazvati egzon 1' jer mu najvjerojatnije u egzon-intron strukturi gena prethodi još jedan egzon. Nukleotidno sravnjenje sljedova ovog pronađenog ispravnog prvog egzona *C. gigas* (92 nt), egzona 1 *M. arenaria* (68 nt) i egzona 1' *M. galloprovincialis* (77 nt) utvrdilo je 42,3% identičnost između ova tri slijeda. Po parovima identičnost je nešto veća, 55,2%. Iako neka mjesta bolje očuvanosti sravnjenih sljedova dijelom odgovaraju očuvanim motivima regulatornih proteina (Pax-4, IRF-8 i ING-4) pronađenih za slijed cDNA3 *M. galloprovincialis* (Tablica 4, Slika 10) tj. egzon 1' (Slika 15C i Slika 16), RegRNA analizom motivi tih regulatornih proteina nisu pronađeni za egzone 1 *M. arenaria* i *C. gigas*. Kod sva tri školjkaša introni nakon ovih prvih egzona (intron 1 *M. arenaria*, intron 1 *C. gigas* i intron 1' *M. galloprovincialis*) imaju istu fazu, fazu 2.



Slika 23. (A) Prikaz regije (26791-26931 nt) genomskog slijeda porodice p53 *C.gigas* (C. gigas gen, JH815916) sa označenim egzonom 1 i tri okvira čitanja (Frame 1, 2, 3). **(B)** Nukleotidno sravnjenje sljedova egzona 1' *M. galloprovincialis* (Mgal1'), egzona 1 *C. gigas* (Cgig1) i egzona 1 *M. arenaria* (Mare1).

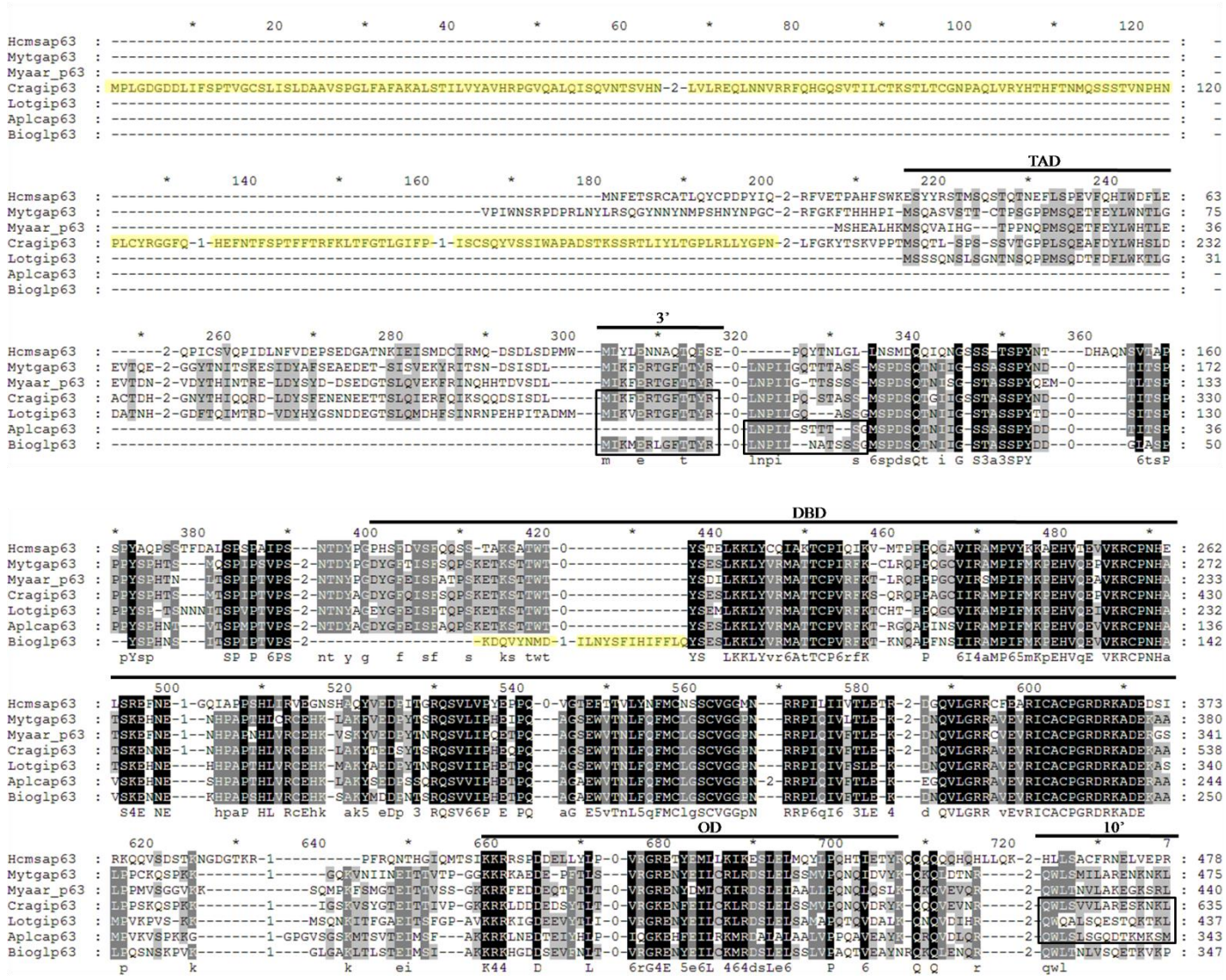
Sravnjenjem slijeda kodirajuće DNA *C. gigas* (AM236465), koja je već identificirana kao TAp63γ ekspresijska izoforma i genomskog slijeda gena porodice p53 *C. gigas* (JH815916), dokazano je postojanje egzona 10' (p63 čovjeka) u strukturi gena porodice p53 *C. gigas* (u postojećoj anotaciji NCBI zapisa egzon 10' nije naznačen). Na ovaj način dokazani egzon 10' gena *C. gigas* pokazuje 55% -tnu identičnost sa egzonom 10' *M. galloprovincialis* (Slika 24).

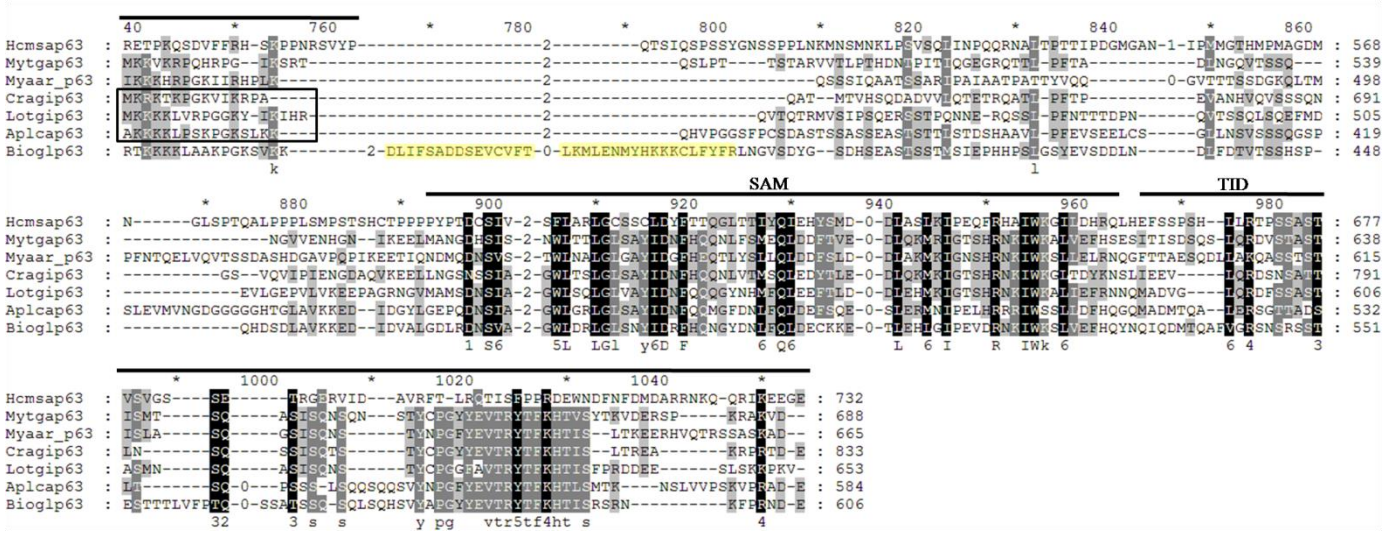
Za razliku od *C. gigas*, N-krajevi *B. glabrata* i *A. californica* značajno su kraći u usporedbi sa izvedenim proteinskim sljedovima *M. galloprovincialis*, *M. arenaria* i referentnim proteinskim slijedom p63 čovjeka (Slika 24). Sravnjenjem proteinskih sljedova utvrđeno je da je početak kodirajućih sljedova proteina *B. glabrata* i *A. californica* unutar *M. galloprovincialis* egzona 4. Pretpostavljenim *B. glabrata* i *A. californica* proteinima nedostaje prvih oko 13 aminokiselina N-kraja, ali one su utvrđene proširenjem okvira čitanja uzvodno od naznačenog prvog egzona na genomskim sljedovima *B. glabrata* i *A. californica* pri čemu su očuvani faza prvog introna uzvodno i AG konsenzus mjesta izrezivanja tog introna (Slika

24). Sravnjenjem proteinskih sljedova je također utvrđeno da su i egzoni 3 i 4 gena porodice p53 *B. glabrata* najvjerojatnije pogrešno anotirani jer ta regija gena kodira za početak DNA-vezneće domene pa se očekuje bolja očuvanost između sljedova predstavnika Mollusca. Produženjem 5' kraja egzona 3 za dodatnih 56 nukleotida uzvodno (pri čemu se unese pomak u okviru čitanja) otkriva se očuvani proteinski slijed (Slika 25). Na sličan način, ali skraćenjem za 38 nukleotida 5' kraja (uz još jedan pomak okvira čitanja koji nadomjesti učinak prvog pomaka okvira čitanja na ostatak otvorenog okvira čitanja), ispravljen je i kodirajući slijed egzona 4 gena *B. glabrata*. Promjenama na egzonima 3 i 4 gena *B. glabrata* očuvane su faze introna 2 i 3 kao i GT/AG konsenzusi mjesta izrezivanja introna. Sravnjenjem *B. glabrata* genomskog slijeda (APKA01029991) sa jedinim poznatim *B. glabrata* „p53-like“ kodirajućim slijedom (GU929337) utvrđeno je da kodirajućem slijedu nedostaju regije koje odgovaraju egzonima 7 i 8 genomskog slijeda. BLASTp pretragom (ograničenom na razred *Mytilus*) utvrđena je sličnost (37%-tna identičnost) kodirajućeg slijeda egzona 7 sa posljednjih 30 aminokiselina *M. galloprovincialis* TAp63γ izoforme (ACD76068) što upućuje na to da *B. glabrata* egzon 7 u organizaciji gena odgovara egzonu 10'. *B. glabrata* egzon 8 BLAST analizom nije pokazao značajnu sličnost sa nijednim poznatim slijedom porodice p53 u GenBank bazi podataka.

Sravnjenjem izvedenih proteinskih sljedova svih predstavnika koljena Mollusca utvrđeno je da je prvih 18 aminokiselina kodiranih sa 5' kraja egzona 9 gena *B. glabrata* pogrešno anotirano. Ako se iz *B. glabrata* proteina izostavi tih 18 aminokiselina zajedno sa aminokiselinama cijelog egzona 8, onda takav ispravljeni egzon 9 slijedi odmah nakon egzona 7, a tu promjenu prati uspostava očuvane faze introna, konsenzusa mjesta izrezivanja te očuvane aminokiseline (glutamina, Q) na mjestu prijelazne aminokiseline između dva egzona (7 i 9). Može se pretpostaviti da *B.glabrata* egzon 8 i prvih 53 nukleotida (18 aminokiselina) egzona 9 ne pripadaju kodirajućem slijedu p63 gena *B. glabrata*.

Kao i za *C.gigas* i *B. glabrata*, egzon 10´ nije anotiran u egzon-intron strukturama genomskih sljedova *A. californica* (NW_004798815) i *L. gigantea* (AMQO01000525). Na osnovu prve


Hcmsap63
Mytgap63
Myaar_p63
Cragip63
Lotgip63
Aplcap63
Bioglp63
TAD
3'
DBD
OD
10'

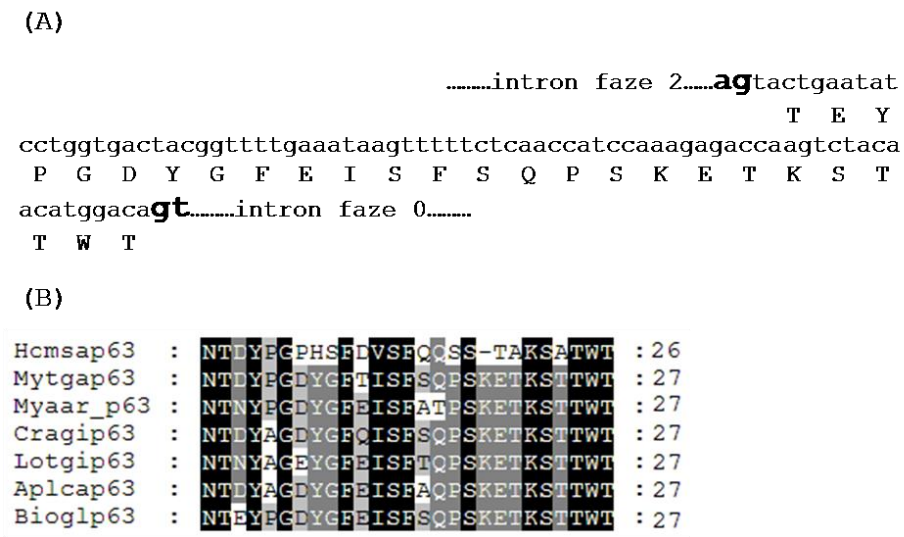


Slika 24. Višestruko sravnjenje izvedenih proteinskih sljedova (sastavljeni proteinski sljedovi; sve izoforme zastupljene) gena p63 čovjeka (*Homo sapiens*- Homsap63) i mekušaca (*Mytilus galloprovincialis*- Mytgap63, *Mya arenaria*- Myaarp63, *Crassostrea gigas*- Cragip63, *Lottia gigantea*- Lotgip63, *Aplysia californica*- Aplcap63, *Biomphalaria glabrata*- Bioglp63). Sravnjenje je dodatno ručno uređeno prema očuvanim regijama proteina. Naznačene su očuvane domene (TA, DBD, OD, SAM i TI) i odabrani egzoni (3' i 10'). Položaji introna naznačeni su brojevima koji odgovaraju fazama introna (0, 1, 2). Aminokiseline koje pretpostavljaju rezultati ovoga rada su crno uokvirene, a one aminokiseline za koje rezultati pretpostavljaju da su pogrešno anotirane su istaknute žuto.

dvije aminokiseline očuvane (100% identičnost) kod egzona 10' *C. gigas*, *B. glabrata* i *M. galloprovincialis*, visoko očuvane prijelazne aminokiseline glutamina (egzoni 10 i 11) i očuvanih nukleotida 3´mjesta izrezivanja introna (AG), a uvažavajući degeneriranost genskog koda, dizajniran je nukleotidni slijed (A-G-A/G-T-G-G-T/C-T) za identifikaciju egzona 10´ na odgovarajućim regijama genomskih sljedova *A. californica* (regija između egzona 7 i 8) i *L. gigantea* (regija između egzona 8 i 9) gdje se egzon 10´očekuje prema sravnjenju izvedenih proteinskih sljedova gena porodice p53 mekušaca. Pronađeno je samo jedno mjesto na oba pretraživana genomska slijeda (*A. californica* i *L. gigantea*) koje odgovara egzonu 10´. Tako je otkriven početak egzona 10´ kod *A. californica* (nt 6761) i *L. gigantea* (nt 12064). Prvi stop kodon nizvodno dodatno je potvrdio pronađene egzone 10´ jer su oba bila očekivae dužine od oko 30 aminokiselina i slabije, ali ipak značajne očuvanosti.

Unatoč visokoj očuvanosti slijeda, egzon 3' nije anotiran na genomskim sljedovima *A. californica*, *L. gigantea*, *B. glabrata* i *C. gigas*. Na osnovu sravnjenja izvedenih proteinskih sljedova i položaja očuvanih domena porodice p53 u analiziranim sljedovima genoma, određeni su introni i/ili genomske regije gdje se očekuje egzon 3'; kod *B. glabrata* i *A. californica* cijela 5´regija uzvodno od prvog anotiranog egzona (4673 nt i 22070 nt duge regije), kod *L. gigantea* intron 2 (6068 nt duga regija) i kod *C. gigas* intron 6 (2156 nt duga regija). Koristeći te odabrane regije kao upit u BLASTx pretraživanju, pronađena je vrlo visoka sličnost (100% identičnost *C. gigas*, 92% *L. gigantea* i 85% za *B. glabrata*) sa prvih 13 aminokiselina (cijela kodirajuća regija egzona 3´) Mytilus ΔNp63/73 sljedova čime je potvrđen 3´egzon kod ova tri analizirana predstavnika koljena Mollusca. Isto pretraživanje nije rezultiralo identifikacijom egzona 3´ kod *A. californica* jer je pretraživana regija diskontinuirana.

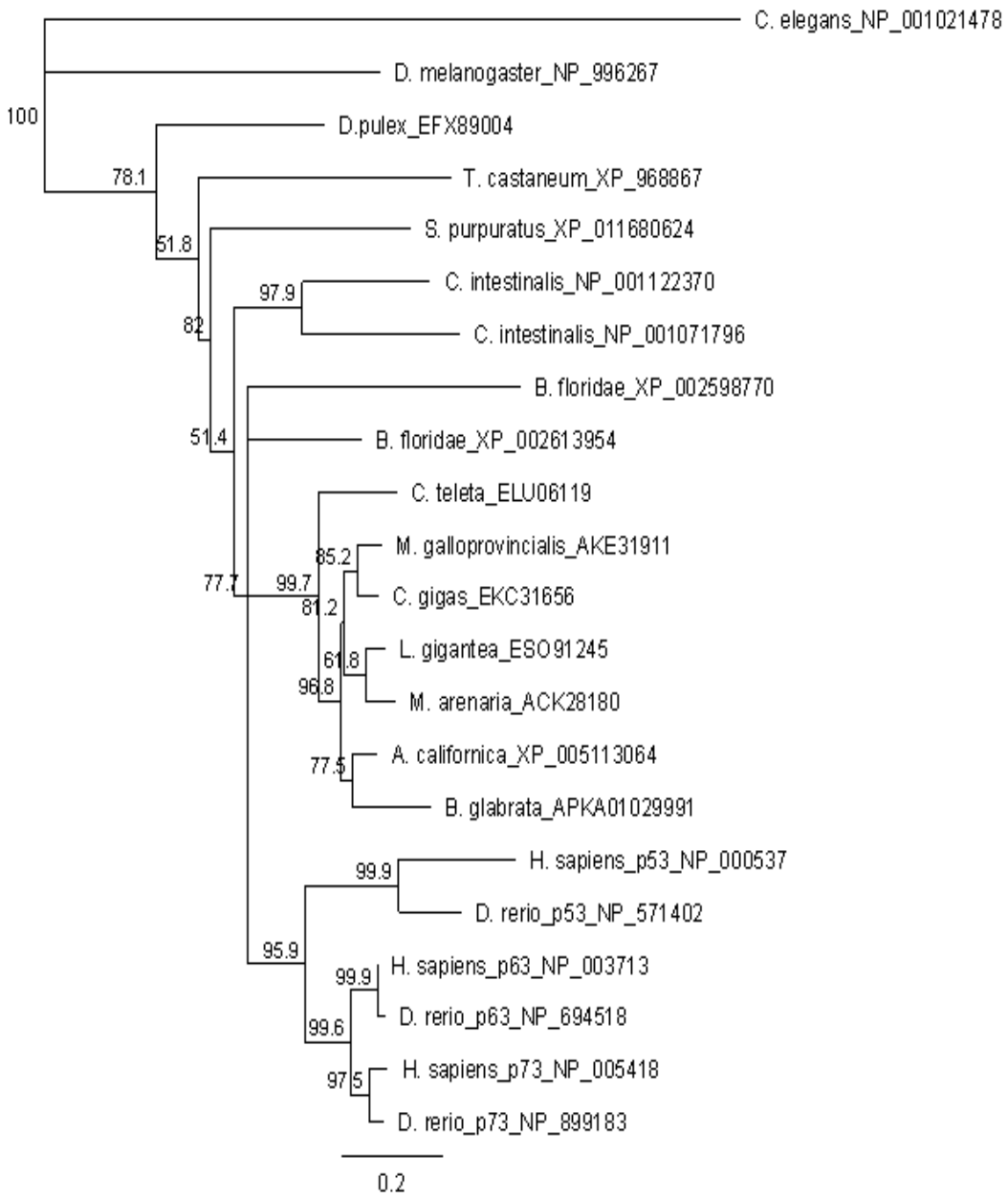
Kao rezultat analize (i ispravaka) genomskih i izvedenih proteinskih sljedova porodice p53 mekušaca *A. californica*, *B. glabrata*, *C. gigas* i *L. gigantea*, 14 mRNA i pripadajućih izvedenih proteina pohranjeno je u GenBank bazu pod pristupnim brojevima TPA: BK009232-BK009245.



Slika 25. Ispravljeni egzon 3 gena p63 puža *Biomphalaria glabrata* i **(B)** sravnjenje izvedenog proteinskog slijeda sa izvedenim proteinskim sljedovima odgovarajućih egzona čovjeka (Homsa, kraj egzona 4) i mekušaca: *Mytilus galloprovincialis* (Mytga, egzon 6), *Mya arenaria* (Myaarp63, početak egzona 6), *Crassostrea gigas* (Cragi, egzon 9), *Lottia gigantea* (Lotgi, egzon 5) i *Aplysia californica* (Apca, egzon 3).

4.7. Filogenetska analiza

Uz proteinske sljedove čovjeka (*H. sapiens*) i šest predstavnika koljena Mollusca (*A. californica*, *B. glabrata*, *C. gigas*, *L. gigantea*, *M. arenaria* i *M. galloprovincialis*) (Tablica 7), za filogenetsku analizu su iz NCBI baze podataka, BLAST pretragom sastavljenih cjelokupnih genoma, preuzeti proteinski sljedovi porodice p53 ribe *Danio rerio*, svitkoglavca *Branchiostoma floridae*, plaštenjaka *Ciona intestinalis*, ježinca *Strongylocentrotus purpuratus*, planktonskog račića *Daphnia pulex*, kukaca *Tribolium castaneum* i *Drosophila melanogaster*, mnogočetinaša *Capitella teleta* i oblića *Caenorhabditis elegans*. U haploidnom genomu ribe *D. rerio* porodica p53 je zastupljena s tri homologa (p53, p63 i p73), u genomima svitkovaca *B. floridae* i *C. intestinalis* s po dva homologa porodice p53. Kod svih ostalih odabranih genoma porodica p53 je zastupljena jednim homologom (Slika 26). Sravnjenje središnjih očuvanih regija proteinskih sljedova (DBD) korištenih za filogenetsku analizu prikazano je u Prilogu 2.



Slika 26. Filogenetsko stablo dobiveno analizom očuvane središnje regije (DBD) sljedova proteina porodice p53(Prilog 2) te primjenom metode sparivanja susjednih grana (neighbour joining, NJ). Kao vanjska grupa (engl. *outgroup*) odabran je slijed *C. elegans*. Pristupni brojevi proteinskih sljedova preuzetih iz GenBank baze prikazani su uz imena vrsta. Dužina grana proporcionalna je broju promjena (bar = 0.2 promjene). Brojevi u čvorištu grana su vrijednosti dobivene metodom samoučitavanja u 1000 ponavljanja, prikazane su samo vrijednosti veće od 50%.

U filogenetskom stablu (Slika 25) vidljivo je da su svi geni kralješnjaka (*D. rerio* i *H. sapiens*) grupirani zajedno i jasno odvojeni u zasebnu granu. Grananje gena p53 dogodilo se ranije nego međusobno odvajanje gena p63 i p73. Svi geni mekušaca također su grupirani zajedno u zasebnu granu, a najbližeg zajedničkog pretka dijele sa genom mnogočetinaša *C. teleta* te tako čine vrlo blisku skupinu uz visoku potporu metodom samoučitavanja (99 %). Ova skupina (mekušci i mnogočetinaši) znatno je bliža kralješnjacima nego svi ostali beskralješnjaci (*D. pulex*, *T. castaneum*, *D. melanogaster*, *C. elegans* i *S.purpuratus*) pa čak i svitkovci (*C. intestinalis* i *B. floridae*) koji imaju po dva gena homologa porodice p53 u svojim haploidnim genomima. Unutar mekušaca uočava se grananje na školjkaše (*M. galloprovincialis* i *C. gigas*) i puževe (*A. californica* i *B. glabrata*) te treću granu (*M. arenaria* i *L. gigantea*) sa predstavnicima oba razreda mekušaca.

5. RASPRAVA

Rezultat analize sljedova cDNA dagnje *M. galloprovincialis* (cDNA1-4) i javno dostupnih sljedova kodirajuće DNA porodice p53 mekušaca, kao što su o tome već raspravljali drugi autori (Goodson i sur., 2006), upućuje na to da je riječ o alternativnim (ekspresijskim) sljedovima koji potječu sa jednog, istog gena. Takva tvrdnja moguća je s obzirom da svi geni porodice p53 mogu eksprimirati alternativne produkte (Murray-Zmijewski i sur., 2006). Proteinska sravnjenja potvrdila su kod sljedova *M. galloprovincialis* domene porodice p53, ranije utvrđene u sljedovima i drugih mekušaca (Kelley i sur., 2001, Jessen-Eller i sur., 2002, Cox i sur., 2003, Muttray i sur., 2005, Goodson i sur., 2006, Muttray i sur., 2007). Rezultati jasno otkrivaju očuvanu središnju regiju gena, gotovo identičnu kod svih proteina iste vrste koljena Mollusca i različite N- i C- krajeve proteina koji nastaju kao rezultat korištenja različitih promotora i alternativnog izrezivanja. Postojanje ΔN fragmenta na N- i SAM domene na C-krajevima proteina upućuje na to da je taj jedan gen porodice p53 mekušaca po svojoj strukturi sličniji p63 i p73, nego p53 čovjeka. S tim u skladu je i sličnost kraćeg C-kraja izvedenih proteinskih sljedova mekušaca sa C-krajem γ izoformi gena p63 čovjeka. Utvrđena egzon-intron struktura cjelovitog slijeda gena porodice p53 kod dagnje *M. galloprovincialis* kao i drugih mekušaca analiziranih u ovom radu, pokazale su najveću sličnost s genom p63 čovjeka. Ipak, gen porodice p53 je kod dagnje, kao predstavnika beskralješnjaka, ishodišni gen porodice te sigurno sadrži strukturalne i funkcionalne osobine sva tri paraloga (p53, p63 i p73). Tako je struktura 5' kraja gena porodice p53 kod dagnje *M. galloprovincialis* osim p63 slična i p73. Iako ima kodirajući egzon 1 kao p63, u strukturi 5' kraja gena porodice p53 kod dagnje *M. galloprovincialis* je utvrđen i nekodirajući egzon 1' te alternativno izrezivanje 5' kraja gena, što su osobine p73 (Kaghad i sur., 1997b, Murray-Zmijewski i sur., 2006). Ispravan prvi egzon *C. gigas* i prvi egzon *M. arenaria* gena porodice p53 upućuju da je ovakva struktura 5' kraja gena porodice p53, kao kod *M. galloprovincialis*, očuvana kod svih školjkaša. Kodirajući slijed *C.gigas* (AM236465) koji je poslužio za identifikaciju egzona 1 u strukturi gena porodice p53 *C. gigas*, opisan je kao cjeloviti slijed dobiven metodom brzog umnažanja 5' i 3' krajeva cDNA (engl. *Rapid Amplification of cDNA ends*, RACE) (Farcy i sur., 2008) pa se očuvani konsenzus izrezivanja introna AG uz 5' kraj egzona 1 može smatrati pouzdanim. Tako je, osim što je nekodirajući, egzon 1 *C. gigas* sličan egzonu 1' *M. galloprovincialis* i po tome što najvjerojatnije u strukturi gena ima ispred sebe još jedan egzon. Iako je poznat genomski slijed (još je oko 20 kb uzvodno od egzona 1 dug genomski slijed *C. gigas*) gdje bi mogli očekivati pronalazak tog još jednog egzona *C. gigas*,

zbog slabe očuvanosti ove regije gena (nema očuvanih domena), sa sigurnošću će ovaj egzon biti utvrđen tek sekvenciranjem ekspresijske izoforme 5' kraja koja započinje tim egzonom. Genomski slijed *M. arenaria* na 5' kraju počinje prvim nukleotidom egzona 1 (cDNA i genomski slijed imaju jednako dug 5' kraj), a još nije poznat sekvencirani genom *M. arenaria* pa za *M. arenaria* nije moguće prema očuvanim konsenzusima izrezivanja introna utvrditi moguće postojanje još jednog egzona ispred egzona 1. Ipak, utvrđena 42% identičnost nukleotidnih sljedova egzona 1 *C. gigas* i *M. arenaria* sa egzonom 1' *M. galloprovincialis* je upućuje da se radi u strukturi gena porodice p53 o homolognim egzonima, a koji odgovaraju egzonu 1 gena p73 čovjeka.

Uz egzon 1', detakcija gena Wrap73 u intronu 1 i „antisense“ organizaciji u odnosu na gen porodice p53 *M. galloprovincialis*, opisuje 5' kraj gena *M. galloprovincialis* kao vrlo sličan 5' kraju p73. Wrap73 detektiran je i u intronima prvih egzona drugih mekušaca (*C.gigas* i *L. gigantea*). Protein Wrap73 pripada porodici proteina sa WD ponavljanjima. Članovi ove porodice pronađeni su kod brojnih eukariota (Smith i sur., 1999), a uključeni su važne stanične procese kao što su stanični ciklus, apoptoza i regulacija gena (Koshizuka i sur., 2001). Sudjelovanje u istim staničnim procesima i genomski smještaj najvjerojatnije upućuje na moguće međusobne interakcije ishodišnih gena Wrap73 i porodice p53 kod dagnje *M. galloprovincialis*, ali to još treba istražiti. Poznato je da gen p53 također ima gen (Wrap53) eksprimiran kao „antisense“ transkript gena p53, a koji ostvaruje interakciju sa 5'UTR mRNA p53 te tako kontrolira indukciju p53 potaknutu oštećenjem DNA (Mahmoudi i sur., 2009). Protein Wrap53 je dobro očuvan među sisavcima i uvijek smješten uz p53 na kromosomu dok su kod svih ne-sisavaca Wrap53 i p53 smješteni na različitim kromosomima. Wrap73 i p73 su kod sisavaca i ne-sisavaca uvijek smješteni jedan uz drugi na istom kromosomu (Mahmouudi i sur., 2009). Uz gen p63 nije poznat gen paralog Wrap73 i Wrap53. Egzon-intron organizacija gena porodice p53 *M. galloprovincialis* upućuje na to da je ishodišni gen porodice p53 homologan genu p63 kralješnjaka/čovjeka, a prisutnost Wrap73 homologa u strukturi 5' kraja može se objasniti da su organizacije genskih lokusa p63 i p73 kralješnjaka/čovjeka rezultat rekombinacijskih događaja tijekom evolucije, a nakon duplikacije koja je odvojila p63 i p73.

Gen porodice p53 *M. galloprovincialis* najveću sličnost sa p53 čovjeka ostvaruje u proteinskom slijedu TA domene (44% identičnost). Ovako visok stupanj identičnosti, s

obzirom da p53 čovjeka sa druga dva paraloga (p63 i p73) ostvaruje identičnost od tek oko 22%, dokaz je strukturalne, ali i signal moguće funkcionalne sličnosti TA domena *M. galloprovincialis* i p53 čovjeka. Za očekivati je da su evolucijski očuvani mehanizam i ciljani geni transaktivacije kao i proteini koji vezanjem na TA reguliraju aktivnost p53 proteina. Očuvana mjesta fosforilacije p53 proteina (Serber i sur., 2002) i vezanja Mdm2 proteina (Chi i sur., 2005) u TA domeni *M. galloprovincialis* to potvrđuju. Osim toga, dokazano je i postojanje funkcionalnog homologa Mdm2 kod školjkaša *Mytilus trossulus* i njegovo vezanje na homologa porodice p53 *M. trossulus* (Muttray i sur., 2010) pa se još pouzdanije može govoriti o ranoj evoluciji ovog mehanizma regulacije porodice p53.

Nekodirajući egzon 1' koji je strukturu ishodišnog gena p53 povezao sa p73 genom, otkriven je pretraživanjem cDNA biblioteke i sekvenciranjem cDNA3 slijeda *M. galloprovincialis*. Pretraživanjem cDNA biblioteke kod *M. galloprovincialis* utvrđena je i ekspresija TAp63γ izoforme (cDNA1 i 2). Djelimični slijed p63α izoforme (cDNA4) upućuje na to da je i ova izoforma C-kraja moguća kod *M. galloprovincialis*. Postojanje ΔN proteinskih izoformi gena p63 *M. galloprovincialis* nije utvrđeno pretraživanjem cDNA biblioteke. Mogući razlog je da su u uzorkovanom tkivu za izradu cDNA biblioteke ove izoforme jednostavno bile slabije ili nikako eksprimirane pa su ih u PCR reakcijama „sakrile“ brojnije, bolje eksprimirane TA izoforme. Poznato je da ekspresija različitih izoformi porodice p53 kod čovjeka ovisi o tipu tkiva, uvjetima u stanici, ali i međusobnim interakcijama članova porodice p53. Ekspresija TAp63 izoformi tako prevladava u diferenciranim, a ΔNp63 u stanicama koje se još dijele (npr. bazalne stanice epitela). ΔN proteinske izoforme inhibiraju aktivnost TA izoformi te na taj način ove dvije izoforme često uzrokuju suprotne učinke u stanici (Nylander i sur., 2002, Petitjean i sur., 2008, Monti i sur., 2014). Dosad ΔN proteinske izoforme nisu ni u literaturi eksperimentalno potvrđene za *M. galloprovincialis*, ali dobivenim genomskim slijedom p63 *M.galloprovincialis* potvrđeno je postojanje egzona 3' u p63 genu *M. galloprovincialis* pa je za očekivati da je ekspresija i ovih izoformi moguća kod *M.galloprovincialis*. Postojanje egzona 3' potvrđeno je i u strukturama gena drugih mekušaca (*M. arenaria, C.gigas, L. gigantea* i *B. glabrata*). Visoki stupanj očuvanosti egzona 3' u koljenu Mollusca upućuje na to da ΔN izoforme imaju evolucijski važnu funkciju, vjerojatno vrlo sličnu funkciji ΔN izoformi kralješnjaka. Ovu pretpostavku treba tek eksperimetalno potvrditi, osobito jer je ekspresija ΔN izoformi kod mekušaca potvrđena samo kod *M. edulis* i *M. trossulus* (Muttray i sur., 2007). Utvrđena struktura gena p63 *M.galloprovincialis* potvrđuje i mogućnost nastanka p63β

izoformi C-kraja iako ni njihova ekspresija nije utvrđena pretraživanjem cDNA biblioteke. Objašnjenje može biti isto kao i za ΔNp63 izoforme jer je poznato da ekspresija i p63β izoformi kod čovjeka ovisi o tipu tkiva, uvjetima u stanici, i međusobnim interakcijama članova porodice p53 (Moll i Slade, 2004, Petitjean i sur., 2008).

Svi mogući alternativni produkti (izoforme) gena porodice p53 dagnje su međusobno različiti proteini te mogu imati različite funkcije što pridonosi dodatnom povećanju mogućnosti ionako brojnih uloga porodice p53. Raznolikost funkcija porodice p53 nameće pitanje koliko su geni ove porodice zapravo dobri kandidati za biomarkere (genotoksičnog) stresa u okolišu. Istraživanja su pokazala da geni porodice p53 mogu biti vrlo dobri indikatori stanja u okolišu (Dondero i sur., 2006, Banni i sur., 2009, Muttray i sur., 2010), ali buduća istraživanja trebaju razjasniti koju od izoformi koristiti za takva mjerenja. Odgovor porodice p53 na onečišćenje u okolišu sigurno neće biti jednostavno povećanje ili smanjenje ekspresije proteina, već mnogo složeniji odgovor koji će ovisiti i o vrsti onečišćenja kao i o nizu drugih faktora koji sudjeluju u ekspresiji gena porodice p53. U prirodnom okolišu sezonski ciklus snažno predodređuje fiziologiju beskralješnjaka, uključujući njihov rast, razmnožavanje i imunost. Sezonske promjene u okolišu utječu tako i na fiziologiju dagnje *Mytilus galloprovincialis* (Petrović i sur., 2004). Fiziološki odgovor na promjene u temperaturi, dostupnost hrane i reproduktivni status kroz različite mjesece, očituje se u promjenama genske ekspresije. Relativna brojnost i prisutnost različitih mRNA odraz su genske ekspresije. Kod mediteranske dagnje je tako u kolovozu zabilježen porast aktivnosti gena vezanih uz imunološki odgovor (matrilin i „p53-like“). Njihova ekspresija vjerojatno je prilagodba na moguće infekcije toksičnim dinoflagelatima koje su najčešće u tom razdoblju godišnjeg ciklusa (Banni i sur., 2011).

Za buduća istraživanja uloga i načina djelovanja gena p53 obitelji kod *M. galloprovincialis* mogu biti korisni pronađeni očuvani motivi vezanja regulatornih proteina u UTR regijama i očekivanim promotorskim regijama sekvenciranog gena porodice p53 *M. galloprovincialis*. Svi regulatorni proteini su evolucijski očuvani pa iako su najbolje opisani kod kralješnjaka, njihovi načini djelovanja su najvjerojatnije očuvani i kod nižih organizama, uključujući i dagnju *M. galloprovincialis*. Tako je za proteine porodice Musashi kod sisavaca potvrđeno da se vežu na 3'UTR mRNA ključnih gena staničnog ciklusa (p21, p27 i p53) te tako sprečavaju translaciju njihovih proteina (Liu i sur., 2014), a kod dagnje su očuvani motivi vezanja ove porodice pronađeni upravo u 3'UTR-ovima sljedova cDNA1 i cDNA2. Kod sisavaca je

poznato da su proteini porodice Musashi eksprimirani tijekom normalnog razvoja živčanih stanica, ali i tijekom tumorigeneze različitih tkiva (Okano i sur., 2002, Nishimoto i Okano, 2010), a to su procesi u koje je uključena i porodica p53, a u procese razvoja i diferencijacije osobito p63 i p73 (posjeduju SAM domenu). Transkripcijski faktori Ets, Pax i Gata porodice reguliraju kod sisavaca također ekspresiju gena (uključujući p53) uključenih u razvoj i diferencijaciju, ali i diobu stanica i apoptozu (Sementchenko i Watson, 2000, Pani i sur., 2002). Pax proteini inhibiraju ekspresiju p53 čovjeka direktnim vezanjem na evolucijski očuvani slijed 5' kraja p53 (netranslatirani prvi egzon), a to je vezno mjesto neophodno i za aktivnost promotora p53 (Stuart i sur., 1995). Kod dagnje *M. galloprovincialis* očuvani motiv vezanja Pax proteina pronađen je upravo na slijedu netranslatiranog egzona 1' pa je to još jedan od dokaza evolucijske očuvanosti ishodišnog gena porodice p53 i sva tri paraloga porodice jasno razdvojena kod kralješnjaka. Mef porodica transkripcijskih faktora su također važni regulatori diferencijacije stanica pa imaju ključnu ulogu u embrionalnom razvoju, osobito mišićnog i živčanog sustava (Mao i sur., 1999). FOXA je transkripcijski faktor koji regulira ekspresiju gena uključenih u stanični ciklus (Williamson i sur., 2005), a za IRF8 je poznato da regulira ekspresiju Mdm2 (Zhou i sur., 2009). Dobro je istražena suradnja proteina ING porodice i p53 u sprečavanju tumorigeneze tako što proteini ING porodice potiču transaktivaciju p53, a ostvaruju i direktnu interakciju sa p53 i Mdm2 (Coles i Jones, 2009, Jafarnejad i Li, 2012). Na molekularnoj razini ING proteini su važan sudionik remodeliranja kromatina (reguliraju aktivnost histon acetiltransferaza i deacetilaza) i regulacije transkripcije ciljanih gena. Sox4 je još jedan važan transkripcijski faktor u razvojnim procesima kralješnjaka te je nužan za diferencijaciju i diobu stanica u mnogim tkivima. Uz to, u mnogim tumorima zabilježena je prekomjerna ekspresija Sox4 (Scharer i sur., 2009). U odgovoru na DNA oštećenje Sox4 potreban je za aktivaciju p53. Sox4 ostvaruje direktnu interakciju sa p53 i tako ga stabilizira jer sprječava Mdm2 posredovanu ubikvitinaciju p53. Osim toga, Sox4 ubrzava acetilaciju p53 olakšavajući stvaranje p300/CBP/p53 kompleksa (Pan i sur., 2009). Zanimljiv je pronalazak veznog mjesta za deltaEF1 protein u regiji očekivanog promotora ΔN izoformi jer je deltaEF1 represor koji na razini transkripcije regulira točno određene članove porodice p53, točnije TAp73, ΔNp73 i ΔNp63, a ne utječe na ekspresiju p53 (Fontemaggi i sur., 2005). Tijekom diferencijacije keratinocita deltaEF1 zauzima regulatorne regije i smanjuje ekspresiju upravo ΔNp63 (De Laurenzi i Melino, 2000). Pronađeno vezno mjesto (p53-RE) p53 porodice u regiji očekivanog promotora ΔN izoformi upućuje na mogućnost da

je i kod dagnje, kao što je dokazano kod čovjeka (Harmes i sur., 2003) ΔN promotor podvrgnut autoregulaciji. Različite izoforme gena p53 dagnje *M. galloprovincialis* će vezanjem na p53-RE vjerojatno značiti i različitu regulaciju ekspresije ΔN izoformi; pozitivnu ili negativnu. Uzimajući u obzir pronađeno vezno mjesto p53, pretpostavljena struktura promotora ΔN izoformi podrazumjeva treći mogući TATA box (TTTATA na 5885-5890 nt) sa dva moguća inicijatora (TTATTTT na 5934-5940 nt i CTATTCC na 5965-5971 nt) kao mjesta početka transkripcije mRNA ΔN izoformi. Tada je 5'UTR ΔN izoformi dužine svega 13 ili 44 nt što bi moglo biti objašnjenje zašto su u bazi poznati *Mytilus* ΔN sljedovi bez 5'UTR tj. započinju ATG kodonom, a opisani su kao cjeloviti sljedovi metodom RACE. Karakterizacija regulatornih elemenata gena kao što je promotor je važan korak u boljem razumjevanju ekspresije i regulacije gena. Zahvaljujući podatcima dostupnim u sekvenciranim genomima, usporedba genomskih sljedova ortologa je postala pouzdan pristup u identifikaciji regulatornih elemenata (Wasserman i Sandelin, 2004). Iako je usporedba promotorske regije ΔNp63 miša sa odgovarajućim regijama kod drugih kralješnjaka (uključujući i čovjeka) otkrila iznimnu identičnost sljedova (oko 97% između miša i čovjeka) u regiji od oko 200 nt uzvodno od TATA-box (Romano i sur., 2006), ista usporedba među sljedovima mekušaca nije otkrila značajnu identičnost među sljedovima.

Utvrđeni polimorfizmi jednog nukleotida (14 razlika) sravnjenjem genomskog slijeda i kodirajućih DNA sljedova gena porodice p53 *M.* galloprovincialis, većinom zahvaćaju treću bazu kodona te se na taj način ne čine kao nasumično unesene razlike metodama PCR-a i sekvenciranja već ove razlike najvjerojatnije predstavljaju alelne varijante gena porodice p53. Ipak, ovi rezultati se mogu smatrati tek preliminarnima jer je svaka razlika (polimorfizam) utvrđena usporedbom samo dva nukleotidna slijeda.

Očuvanost strukture analiziranih gena porodice p53 mekušaca prati i očuvanost položaja i faze introna. Pozicije introna u ortolozima različitih vrsta su najčešće vrlo dobro očuvane te mogu poslužiti kao filogenetski marker. Pri tome udio očuvanih intronskih pozicija ortologa opada što je veća evolucijska udaljenost među vrstama (Rogozin i sur., 2005). Introni imaju vrlo sporu stopu ugradnje ili gubitka iz genske strukture te na taj način zadržavaju „filogenetski signal“ duže od kodirajućih sljedova. Osim toga, gubitak introna na određenoj poziciji je nereverzibilan događaj pa je izbjegnut problem povratnih mutacija kao u kodirajućim sljedovima (Patthy, 1999, Rogozin i sur., 2003). Ipak, sravnjenje intronskih

pozicija uvelike ovisi o očuvanosti aminokiselinskog slijeda te je osobito nepouzdano sravnjenje u regijama slabe očuvanosti. Ako intron zadrži svoju poziciju u genu, onda je i faza introna značajno evolucijski očuvano svojstvo gena (Rogozin i sur., 2012).

Iako je posljednjih godina došlo do značajnog povećanja broja i raznolikosti sekvenciranih genoma, današnji pristupi detekcije gena u sekvenciranim genomima najčešće ne detektiraju egzone koji kodiraju transkripcijske varijante alternativnog izrezivanja. Alternativno izrezivanje introna je jedan od glavnih načina povećanja složenosti transkriptoma. Više od 86% gena čovjeka ima alternativno izrezivanje introna, a prosječan broj transkripcijskih varijanti po lokusu je 5,7 (Harrow i sur., 2006). Sve pogreške u anotaciji pronađene u analiziranim genomskim sljedovima mekušaca potvrđuju da se unatoč razvoju automatskog (računalnog) pronalaska gena u sekvenciranom genomu, najpouzdanije i najtočnije anotacije gena i dalje postižu sravnjenjima cDNA i proteina na sljedove genoma i utvrđivanjem granica egzona i transkripata (Harrow i sur., 2009).

Izgled filogenetskog stabla i broj gena porodice p53 u haploidnim genomima analiziranih mekušaca i drugih organizama, upućuju na grananje jednog, ishodišnog gena u tri člana porodice samo kod kralješnjaka, a to je u skladu s rezultatima drugih autora (Yang i sur., 2002, Belyi i sur., 2010). Dva člana porodice p53 kod predstavnika svitkovaca (*C. intestinalis* i *B. floridae*) te jedan kod ježinca *S. purpuratus* i ostalih anliziranih beskralješnjaka, upućuju na to da se prva duplikacija ishodišnog gena porodice p53 nije dogodila tek kod riba hrskavičnjača, nego i ranije. Ipak, zanimljivo je da su ti duplicirani geni u filogenetskom stablu udaljeniji od svojih homologa kod kralješnjaka, nego što je to gen porodice p53 *M. galloprovincialis* i ostalih mekušaca. Za potpuniju analizu evolucije porodice p53 potrebano je usporediti i strukture gena te obuhvatiti što veći broj različitih organizama. Važno je i bolje poznavanje funkcije porodice kod beskralješnjaka. Bliska povezanost gena porodice p53 mnogočetinaša *C. teleta* sa genima svih mekušaca u skladu je sa novijim istraživanjima koja pokazuju da su organizacije genoma, strukture i funkcije gena ove dvije skupine životinja sličnije genomima deuterostomičnih beskralješnjaka (npr. svitkoglavaca i ježinaca), nego drugih protostomičnih kao što su kukci ili oblići (Simakov i sur., 2013). Strukturalna i funkcionalna očuvanost ishodišnog gena porodice p53 kod mekušaca svrstava ih u korisne, dobro istražene modelne organizme istraživanja u biomedicini i/ili biomonitoringu.

6. ZAKLJUČCI

Mediteranska dagnja *Mytilus galloprovincialis*, kao i svi ostali analizirani predstavnici koljena Mollusca (*Aplysia californica*, *Biomphalaria glabrata*, *Crassostrea gigas* i *Lottia gigantea*) u haploidnom genomu ima samo jedan gen porodice p53, prisutan u jednoj kopiji.

Ukupna egzon-intron organizacija tog jednog, ishodišnog gena, vrlo je slična genu p63 čovjeka pa se u nomenkaturi gena porodice p53 kod mediteranske dagnje *Mytilus galloprovincialis* i ostalih analiziranih predstavnika koljena Mollusca (*A. californica*, *B. glabrata*, *C. gigas* i *L. gigantea*) može koristiti samo naziv p63.

Svi kodirajući sljedovi *M. galloprovincialis* (cDNA1-4) kodiraju za TAp63γ ili TAp63α proteinske izoforme gena p63, a i svi ostali poznati sljedovi porodice p53 u GenBank bazi također kodiraju za ekspresijske izoforme gena p63.

Struktura gena p63 očuvana je među predstavnicima koljena Mollusca. Očuvana struktura gena može olakšati anotaciju, funkcionalne i evolucijske analize nukleotidnog zapisa bilo kojeg novog ortologa.

Analize sravnjenja cDNA s genomskim sljedovima značajno pridonose cjelovitoj i pouzdanoj anotaciji gena.

Anotacije gena p63 u sekvenciranim genomima različito su pouzdane. Prema našoj analizi, genom *L. gigantea* ima najtočnije anotiran gen p63. Osim egzona 3´i 10´, nisu pronađena druga neslaganja ili nedostatci u anotaciji u odnosu na strukturu gena predviđenu usporedbom sa drugim p63 sljedovima mekušaca.

7. LITERATURA

Altschul, SF, Gish, W, Miller, W, Myers, EW, Lipman, DJ (1990) Basic local alignment search tool. Journal of Molecular Biology. 215:403-410.

Baeza Garcia, A, Pierce, RJ, Gourbal, B, Werkmeister, E, Colinet, D, Reichhart, J-M, Dissous, C, Coustau, C (2010) Involvement of the Cytokine MIF in the Snail Host Immune Response to the Parasite *Schistosoma mansoni*. PLoS Pathogens. 6:e1001115.

Banni, M, Negri, A, Mignone, F, Boussetta, H, Viarengo, A, Dondero, F (2011) Gene expression rhythms in the mussel *Mytilus galloprovincialis* (Lam.) across an annual cycle. PLoS ONE. 6:

Banni, M, Negri, A, Rebelo, M, Rapallo, F, Boussetta, H, Viarengo, A, Dondero, F (2009) Expression analysis of the molluscan p53 protein family mRNA in mussels (*Mytilus* spp.) exposed to organic contaminants. Comparative Biochemistry and Physiology-Part C Toxicology Pharmacology. 149:414-418.

Barber, BJ (2004) Neoplastic diseases of commercially important marine bivalves. Aquatic Living Resources. 17:449-466.

Belyi, VA, Ak, P, Markert, E, Wang, H, Hu, W, Puzio-Kuter, A, Levine, AJ (2010) The origins and evolution of the p53 family of genes. Cold Spring Harbor Perspectives in Biology. 2:

Bénard, J, Douc-Rasy, S, Ahomadegbe, J-C (2003) TP53 family members and human cancers. Human Mutation. 21:182-191.

Benson, G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Research. 27:573-580.

Berriman, M, Haas, BJ, LoVerde, PT, Wilson, RA, Dillon, GP i sur., (2009) The genome of the blood fluke *Schistosoma mansoni*. Nature. 460:352-358.

Bourdon, JC, Fernandes, K, Murray-Zmijewski, F, Liu, G, Diot, A, Xirodimas, DP, Saville, MK, Lane, DP (2005) p53 isoforms can regulate p53 transcriptional activity. Genes and Development. 19:2122-2137.

Brandt, T, Petrovich, M, Joerger, AC, Veprintsev, DB (2009) Conservation of DNA-binding specificity and oligomerisation properties within the p53 family. BMC Genomics. 10:

Brodsky, MH, Weinert, BT, Tsang, G, Rong, YS, McGinnis, NM, Golic, KG, Rio, DC, Rubin, GM (2004) Drosophila melanogaster MNK/Chk2 and p53 Regulate Multiple DNA Repair and Apoptotic Pathways following DNA Damage. Molecular and Cellular Biology. 24:1219-1231.

Chang, T-H, Huang, H-Y, Hsu, J, Weng, S-L, Horng, J-T, Huang, H-D (2013) An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. BMC Bioinformatics. 14:S4.

Chi, S-W, Lee, S-H, Kim, D-H, Ahn, M-J, Kim, J-S, Woo, J-Y, Torizawa, T, Kainosho, M, Han, K-H (2005) Structural Details on mdm2-p53 Interaction. Journal of Biological Chemistry. 280:38795-38802.

Ciocan, CM, Rotchell, JM (2005) Conservation of cancer genes in the marine invertebrate *Mytilus edulis*. Environmental Science & Technology. 39:3029-3033.

Coles, AH, Jones, SN (2009) The ING Gene Family in the Regulation of Cell Growth and Tumorigenesis. Journal of cellular physiology. 218:45-57.

Collavin, L, Lunardi, A, Del Sal, G (2010) p53-family proteins and their regulators: hubs and spokes in tumor suppression. Cell Death & Differentiation. 17:901-911.

Cox, RL, Stephens, RE, Reinisch, CL (2003) p63/73 homologues in surf clam: novel signaling motifs and implications for control of expression. Gene. 320:49-58.

D'Erchia, AM, Tullo, A, Pesole, G, Saccone, C, Sbisà, E (2003) p53 gene family: Structural, functional and evolutionary features. Current Genomics. 4:13-26.

Danilova, N, Sakamoto, KM, Lin, S (2008) p53 family in development. Mechanisms of Development. 125:919-931.

De Laurenzi, V, Melino, G (2000) Evolution of functions within the p53/p63/p73 family. Annals of the New York Academy of Sciences. 926:90-100.

Derry, WB, Putzke, AP, Rothman, JH (2001) Caenorhabditis elegans p53: Role in apoptosis, meiosis, and stress resistance. Science. 294:591-595.

Derry, WB, Bierings, R, van Iersel, M, Satkunendran, T, Reinke, V, Rothman, JH (2006) Regulation of developmental rate and germ cell proliferation in Caenorhabditis elegans by the p53 gene network. Cell Death & Differentiation. 14:662-670.

DeYoung, MP, Ellisen, LW (2007) p63 and p73 in human cancer: defining the network. Oncogene. 26:5169-5183.

Dondero, F, Piacentini, L, Marsano, F, Rebelo, M, Vergani, L, Venier, P, Viarengo, A (2006) Gene transcription profiling in pollutant exposed mussels (*Mytilus* spp.) using a new low-density oligonucleotide microarray. Gene. 376:24-36.

Dotsch, V, Bernassola, F, Coutandin, D, Candi, E, Melino, G (2010) p63 and p73, the Ancestors of p53. Cold Spring Harbor Perspectives in Biology. 2:14.

Farcy, E, Fleury, C, Lelong, C, Dubos, MP, Voiseux, C, Fiévet, B, Lebel, JM (2008) Molecular cloning of a new member of the p53 family from the Pacific oyster Crassostrea gigas and seasonal pattern of its transcriptional expression level. Marine Environmental Research. 66:300-308.

Fontemaggi, G, Gurtner, A, Damalas, A, Costanzo, A, Higashi, Y, Sacchi, A, Strano, S, Piaggio, G, Blandino, G (2005) [delta]EF1 repressor controls selectively p53 family members during differentiation. Oncogene. 24:7273-7280.

Gagné, F, Burgeot, T, Hellou, J, St-Jean, S, Farcy, É, Blaise, C (2008) Spatial variations in biomarkers of Mytilus edulis mussels at four polluted regions spanning the Northern Hemisphere. Environmental Research. 107:201-217.

Goodson, MS, Crookes-Goodson, WJ, Kimbell, JR, McFall-Ngai, MJ (2006) Characterization and Role of p53 Family Members in the Symbiont-Induced Morphogenesis of the Euprymna scolopes Light Organ. The Biological Bulletin. 211:7-17.

Gosling, E (1992) The mussel Mytilus: ecology, physiology, genetics and culture. Elsevier,

Hamer, B, Hamer, DP, Müller, WEG, Batel, R (2004) Stress-70 proteins in marine mussel Mytilus galloprovincialis as biomarkers of environmental pollution: A field study. Environment International. 30:873-882.

Hamer, B, Jakšić, Z, Pavičić-Hamer, D, Perić, L, Medaković, D, Ivanković, D, Pavičić, J, Zilberberg, C, Schröder, HC, Müller, WEG, Smodlaka, N, Batel, R (2008) Effect of hypoosmotic stress by low salinity acclimation of Mediterranean mussels Mytilus galloprovincialis on biological parameters used for pollution assessment. Aquatic Toxicology. 89:137-151.

Harmes, DC, Bresnick, E, Lubin, EA, Watson, JK, Heim, KE, Curtin, JC, Suskind, AM, Lamb, J, DiRenzo, J (2003) Positive and negative regulation of [Delta]N-p63 promoter activity by p53 and [Delta]N-p63-[alpha] contributes to differential regulation of p53 target genes. Oncogene. 22:7607-7616.

Harrow, J, Nagy, A, Reymond, A, Alioto, T, Patthy, L, Antonarakis, S, Guigo, R (2009) Identifying protein-coding genes in genomic sequences. Genome Biology. 10:201.

Harrow, J, Denoeud, F, Frankish, A, Reymond, A, Chen, C-K, Chrast, J, Lagarde, J, Gilbert, J, Storey, R, Swarbreck, D, Rossier, C, Ucla, C, Hubbard, T, Antonarakis, S, Guigo, R (2006) GENCODE: producing a reference annotation for ENCODE. Genome Biology. 7:S4.

Holbrook, LAC, Butler, RA, Cashon, RE, Van Beneden, RJ (2009) Soft-shell clam (Mya arenaria) p53: A structural and functional comparison to human p53. Gene. 433:81-87.

Hollstein, M, Shomer, B, Greenblatt, M, Soussi, T, Hovig, E, Montesano, R, Harris, CC (1996) Somatic point mutations in the p53 gene of human tumors and cell lines: Updated compilation. Nucleic Acids Research. 24:137-140.

Hollstein, M, Rice, K, Greenblatt, MS, Soussi, T, Fuchs, R, Sørlie, T, Hovig, E, Smith-Sørensen, B, Montesano, R, Harris, CC (1994) Database of p53 gene somatic mutations in human tumors and cell lines. Nucleic Acids Research. 22:3551-3555.

Holt, RA, Subramanian, GM, Halpern, A, Sutton, GG, Charlab, R i sur., (2002) The genome sequence of the malaria mosquito Anopheles gambiae. Science. 298:129-149.

Irwin, MS, Kaelin, WG (2001) p53 Family Update: p73 and p63 Develop Their Own Identities. Cell Growth Differentiation. 12:337-349.

Isobe, M, Emanuel, BS, Givol, D, Oren, M, Croce, CM (1986) Localization of gene for human p53 tumour antigen to band 17p13. Nature. 320:84-85.

Jafarnejad, S, Li, G (2012) Regulation of p53 by ING family members in suppression of tumor initiation and progression. Cancer and Metastasis Reviews. 31:55-73.

Jarque, S, Prats, E, Olivares, A, Casado, M, Ramón, M, Piña, B (2014) Seasonal variations of gene expression biomarkers in Mytilus galloprovincialis cultured populations: Temperature, oxidative stress and reproductive cycle as major modulators. Science of the Total Environment. 499:363-372.

Jessen-Eller, K, Kreiling, JA, Begley, GS, Steele, ME, Walker, CW, Stephens, RE, Reinisch, CL (2002) A new invertebrate member of the p53 gene family is developmentally expressed and responds to polychlorinated biphenyls. Environmental Health Perspectives. 110:377-385.

Joerger, AC, Rajagopalan, S, Natan, E, Veprintsev, DB, Robinson, CV, Fersht, AR (2009) Structural evolution of p53, p63, and p73: Implication for heterotetramer formation. Proceedings of the National Academy of Sciences of the United States of America. 106:17705-17710.

Kaghad, M, Bonnet, H, Yang, A, Creancier, L, Biscan, JC, Valent, A, Minty, A, Chalon, P, Lelias, JM, Dumont, X, Ferrara, P, McKeon, F, Caput, D (1997a) Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. Cell. 90:809-819.

Kaghad, M, Bonnet, H, Yang, A, Creancier, L, Biscan, J-C, Valent, A, Minty, A, Chalon, P, Lelias, J-M, Dumont, X, Ferrara, P, McKeon, F, Caput, D (1997b) Monoallelically Expressed Gene Related to p53 at 1p36, a Region Frequently Deleted in Neuroblastoma and Other Human Cancers. Cell. 90:809-819.

Kapustin, Y, Souvorov, A, Tatusova, T, Lipman, D (2008) Splign: algorithms for computing spliced alignments with identification of paralogs. Biology Direct. 3:20.

Kelley, ML, Winge, P, Heaney, JD, Stephens, RE, Farell, JH, Van Beneden, RJ, Reinisch, CL, Lesser, MP, Walker, CW (2001) Expression of homologues for p53 and p73 in the softshell clam (Mya arenaria), a naturally-occurring model for human cancer. Oncogene. 20:748-758.

Kim, CA, Bowie, JU (2003) SAM domains: Uniform structure, diversity of function. Trends in Biochemical Sciences. 28:625-628.

King, N, Westbrook, MJ, Young, SL, Kuo, A, Abedin, M i sur., (2008) The genome of the choanoflagellate Monosiga brevicollis and the origin of metazoans. Nature. 451:783-788.

Koshizuka, Y, Ikegawa, S, Sano, M, Nakamura, K, Nakamura, Y (2001) Isolation, characterization, and mapping of the mouse and human WDR8 genes, members of a novel WD-repeat gene family. Genomics. 72:252-259.

Kozak, M (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Research. 15:8125-8148.

Levrero, M, De Laurenzi, V, Costanzo, A, Gong, J, Wang, JY, Melino, G (2000) The p53/p63/p73 family of transcription factors: overlapping and distinct functions. Journal of Cell Science. 113 (Pt 10):1661-1670.

Liu, X, Yang, WT, Zheng, PS (2014) Msi1 promotes tumor growth and cell proliferation by targeting cell cycle checkpoint proteins p21, p27 and p53 in cervical carcinomas. Oncotarget. 5:10870-10885.

Mahmoudi, S, Henriksson, S, Corcoran, M, Méndez-Vidal, C, Wiman, KG, Farnebo, M (2009) Wrap53, a Natural p53 Antisense Transcript Required for p53 Induction upon DNA Damage. Molecular Cell. 33:462-471.

Maisse, C, Guerrieri, P, Melino, G (2003) p73 and p63 protein stability: the way to regulate function? Biochemical Pharmacology. 66:1555-1561.

Mangiulli, M, Valletti, A, Caratozzolo, MF, Tullo, A, Sbisa, E, Pesole, G, D'Erchia, AM (2009) Identification and functional characterization of two new transcriptional variants of the human p63 gene. Nucleic Acids Research. 37:6092-6104.

Mao, Z, Bonni, A, Xia, F, Nadal-Vicens, M, Greenberg, ME (1999) Neuronal Activity-Dependent Cell Survival Mediated by Transcription Factor MEF2. Science. 286:785-790.

Marchler-Bauer, A, Lu, S, Anderson, JB, Chitsaz, F, Derbyshire, MK, DeWeese-Scott, C, Fong, JH, Geer, LY, Geer, RC, Gonzales, NR, Gwadz, M, Hurwitz, DI, Jackson, JD, Ke, Z, Lanczycki, CJ, Lu, F, Marchler, GH, Mullokandov, M, Omelchenko, MV, Robertson, CL,

Song, JS, Thanki, N, Yamashita, RA, Zhang, D, Zhang, N, Zheng, C, Bryant, SH (2011) CDD: A Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Research. 39:D225-D229.

Matoničkin, I, Habdija, I, Primc-Habdija, B (1998) Beskralješnjaci- biologija nižih avertebrata. Školska knjiga, Zagreb.

Mavinahalli, JN, Madhumalar, A, Beuerman, RW, Lane, DP, Verma, C (2010) Differences in the transactivation domains of p53 family members: A computational study. BMC Genomics. 11:

Melino, G, Lu, X, Gasco, M, Crook, T, Knight, RA (2003) Functional regulation of p73 and p63: Development and cancer. Trends in Biochemical Sciences. 28:663-670.

Moll, UM, Slade, N (2004) p63 and p73: Roles in Development and Tumor Formation11National Cancer Institute. Molecular Cancer Research. 2:371-386.

Moll, UM, Erster, S, Zaika, A (2001) p53, p63 and p73 - Solos, alliances and feuds among family members. Biochimica et Biophysica Acta - Reviews on Cancer. 1552:47-59.

Monti, P, Ciribilli, Y, Bisio, A, Foggetti, G, Raimondi, I, Campomenosi, P, Menichini, P, Fronza, G, Inga, A (2014) ΔN-P63α and TA-P63α exhibit intrinsic differences in transactivation specificities that depend on distinct features of DNA target sites. Oncotarget. 5:2116-2130.

Mosner, J, Mummenbrauer, T, Bauer, C, Sczakiel, G, Grosse, F, Deppert, W (1995) Negative feedback regulation of wild-type p53 biosynthesis. EMBO Journal. 14:4442-4449.

Murray-Zmijewski, F, Lane, DP, Bourdon, JC (2006) p53//p63//p73 isoforms: an orchestra of isoforms to harmonise cell differentiation and response to stress. Cell Death Differ. 13:962-972.

Muttray, AF, Cox, RL, Reinisch, CL, Baldwin, SA (2007) Identification of DeltaN isoform and polyadenylation site choice variants in Molluscan p63/p73-like homologues. Marine Biotechnology. 9:217-230.

Muttray, AF, O'Toole, TF, Morrill, W, Van Beneden, RJ, Baldwin, SA (2010) An invertebrate mdm homolog interacts with p53 and is differentially expressed together with p53 and ras in neoplastic Mytilus trossulus haemocytes. Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology. 156:298-308.

Muttray, AF, Cox, RL, St-Jean, S, van Poppelen, P, Reinisch, CL, Baldwin, SA (2005) Identification and phylogenetic comparison of p53 in two distinct mussel species (Mytilus). Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology. 140:237-250.

Nedelcu, AM, Tan, C (2007) Early diversification and complex evolutionary history of the p53 tumor suppressor gene family. Development Genes and Evolution. 217:801-806.

Nicholas, KBN, H. B. Jr.; Deerfield, D. W. II. (1997) GeneDoc: Analysis and Visualization of Genetic Variation EMBNEW News. 4:

Nishimoto, Y, Okano, H (2010) New insight into cancer therapeutics: Induction of differentiation by regulating the Musashi/Numb/Notch pathway. Cell Research. 20:1083-1085.

Nylander, K, Vojtesek, B, Nenutil, R, Lindgren, B, Roos, G, Zhanxiang, W, Sjostrom, B, Dahlqvist, A, Coates, PJ (2002) Differential expression of p63 isoforms in normal tissues and neoplastic cells. J Pathol. 198:417-427.

Okano, H, Imai, T, Okabe, M (2002) Musashi: a translational regulator of cell fate. Journal of Cell Science. 115:1355-1359.

Ollmann, M, Young, LM, Di Como, CJ, Karim, F, Belvin, M, Robertson, S, Whittaker, K, Demsky, M, Fisher, WW, Buchman, A, Duyk, G, Friedman, L, Prives, C, Kopczynski, C (2000) Drosophila p53 is a structural and functional homolog of the tumor suppressor p53. Cell. 101:91-101.

Oren, M (2003) Decision making by p53: life, death and cancer. Cell Death & Differentiation. 10:431-442.

Pan, X, Zhao, J, Zhang, WN, Li, HY, Mu, R, Zhou, T, Zhang, HY, Gong, WL, Yu, M, Man, JH, Zhang, PJ, Li, AL, Zhang, XM (2009) Induction of SOX4 by DNA damage is critical for p53 stabilization and function. Proceedings of the National Academy of Sciences of the United States of America. 106:3788-3793.

Pani, L, Horal, M, Loeken, MR (2002) Rescue of neural tube defects in Pax-3-deficient embryos by p53 loss of function: implications for Pax-3- dependent development and tumorigenesis. Genes & Development. 16:676-680.

Pankow, S, Bamberger, C (2007) The p53 tumor suppressor-like protein nvp63 mediates selective germ cell death in the sea anemone Nematostella vectensis. PLoS ONE. 2:

Patthy, L (1999) Genome evolution and the evolution of exon-shuffling - A review. Gene. 238:103-114.

Petitjean, A, Hainaut, P, Caron de Fromentel, C (2006) TP63 gene in stress response and carcinogenesis: a broader role than expected. Bulletin du cancer. 93:E126-135.

Petitjean, A, Ruptier, C, Tribollet, V, Hautefeuille, A, Chardon, F, Cavard, C, Puisieux, A, Hainaut, P, Caron de Fromentel, C (2008) Properties of the six isoforms of p63: p53-like regulation in response to genotoxic stress and cross talk with ΔNp73. Carcinogenesis. 29:273-281.

Petrović, S, Semenčić, L, Ozretić, B, Ozretić, M (2004) Seasonal variations of physiological and cellular biomarkers and their use in the biomonitoring of north adriatic coastal waters (Croatia). Marine Pollution Bulletin. 49:713-720.

Pietsch, EC, Sykes, SM, McMahon, SB, Murphy, ME (2008) The p53 family and programmed cell death. Oncogene. 27:6507-6521.

Putnam, NH, Srivastava, M, Hellsten, U, Dirks, B, Chapman, J, Salamov, A, Terry, A, Shapiro, H, Lindquist, E, Kapitonov, VV, Jurka, J, Genikhovich, G, Grigoriev, IV, Lucas, SM, Steele, RE, Finnerty, JR, Technau, U, Martindale, MQ, Rokhsar, DS (2007) Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science. 317:86-94.

Putnam, NH, Butts, T, Ferrier, DEK, Furlong, RF, Hellsten, U i sur., (2008) The amphioxus genome and the evolution of the chordate karyotype. Nature. 453:1064-1071.

Qian, Y, Chen, X (2013) Senescence regulation by the p53 protein family. Methods in molecular biology (Clifton, N.J.). 965:37-61.

Rogozin, IB, Sverdlov, AV, Babenko, VN, Koonin, EV (2005) Analysis of evolution of exon-intron structure of eukaryotic genes. Briefings in Bioinformatics. 6:118-134.

Rogozin, IB, Carmel, L, Csuros, M, Koonin, EV (2012) Origin and evolution of spliceosomal introns. Biology Direct. 7:11.

Rogozin, IB, Babenko, VN, Fedorova, ND, Jackson, JD, Jacobs, AR, Krylov, DM, Makarova, KS, Mazumder, R, Mekhedov, SL, Mirkin, BG, Nikolskaya, AN, Rao, BS, Smirnov, S, Sorokin, AV, Sverdlov, AV, Vasudevan, S, Wolf, YI, Yin, JJ, Natale, DA, Koonin, EV. (2003) Evolution of eukaryotic gene repertoire and gene structure: Discovering the unexpected dynamics of genome evolution. Pages 293-301. Cold Spring Harbor Symposia on Quantitative Biology.

Romano, R-A, Birkaya, B, Sinha, S (2006) Defining the Regulatory Elements in the Proximal Promoter of [Delta]Np63 in Keratinocytes: Potential Roles for Sp1/Sp3, NF-Y, and p63. Journal of Investigative Dermatology. 126:1469-1479.

Ruiz, M, López, C (2013) Neoplasms in clams. Nova Science Publishers, Inc., p. 149-162.

Rutkowski, R, Hofmann, K, Gartner, A (2010) Phylogeny and function of the invertebrate p53 superfamily. Cold Spring Harbor Perspectives in Biology. 2:a001131.

Saavedra, C, Bachère, E (2006) Bivalve genomics. Aquaculture. 256:1-14.

Scharer, CD, McCabe, CD, Ali-Seyed, M, Berger, MF, Bulyk, ML, Moreno, CS (2009) Genome-Wide Promoter Analysis of the SOX4 Transcriptional Network in Prostate Cancer Cells. Cancer Research. 69:709-717.

Schumacher, B, Hofmann, K, Boulton, S, Gartner, A (2001) The C. elegans homolog of the p53 tumor suppressor is required for DNA damage-induced apoptosis. Current Biology. 11:1722-1727.

Schumacher, B, Schertel, C, Wittenburg, N, Tuck, S, Mitani, S, Gartner, A, Conradt, B, Shaham, S (2004) C. elegans ced-13 can promote apoptosis and is induced in response to DNA damage. Cell Death & Differentiation. 12:153-161.

Scoumanne, A, Harms, KL, Chen, X (2005) Structural basis for gene activation by p53 family members. Cancer Biology and Therapy. 4:1178-1185.

Sementchenko, VI, Watson, DK (2000) Ets target genes: past, present and future. Oncogene. 19:6533.

Serber, Z, Lai, HC, Yang, A, Ou, HD, Sigal, MS, Kelly, AE, Darimont, BD, Duijf, PHG, van Bokhoven, H, McKeon, F, Dotsch, V (2002) A C-terminal inhibitory domain controls the activity of p63 by an intramolecular mechanism. Molecular and Cellular Biology. 22:8601-8611.

Simakov, O, Marletaz, F, Cho, SJ, Edsinger-Gonzales, E, Havlak, P, Hellsten, U, Kuo, DH, Larsson, T, Lv, J, Arendt, D, Savage, R, Osoegawa, K, de Jong, P, Grimwood, J, Chapman, JA, Shapiro, H, Aerts, A, Otillar, RP, Terry, AY, Boore, JL, Grigoriev, IV, Lindberg, DR, Seaver, EC, Weisblat, DA, Putnam, NH, Rokhsar, DS (2013) Insights into bilaterian evolution from three spiralian genomes. Nature. 493:526-531.

Smith, TF, Gaitatzes, C, Saxena, K, Neer, EJ (1999) The WD repeat: a common architecture for diverse functions. Trends in Biochemical Sciences. 24:181-185.

St.-Jean, SD, Stephens, RE, Courtenay, SC, Reinisch, CL (2005) Detecting p53 family proteins in haemocytic leukemia cells of Mytilus edulis from Pictou Harbour, Nova Scotia, Canada. Canadian Journal of Fisheries and Aquatic Sciences. 62:2055-2066.

Stephens, RE, Walker, CW, Reinisch, CL (2001) Multiple protein differences distinguish clam leukemia cells from normal hemocytes: evidence for the involvement of p53 homologues. Comparative Biochemistry and Physiology. Toxicology & Pharmacology. 129:329-338.

Strano, S, Rossi, M, Fontemaggi, G, Munarriz, E, Soddu, S, Sacchi, A, Blandino, G (2001) From p63 to p53 across p73. FEBS Letters. 490:163-170.

Stuart, ET, Haffner, R, Oren, M, Gruss, P (1995) Loss of p53 function through PAX-mediated transcriptional repression. The EMBO Journal. 14:5638-5645.

Suh, EK, Yang, A, Kettenbach, A, Bamberger, C, Michaelis, AH, Zhu, Z, Elvin, JA, Bronson, RT, Crum, CP, McKeon, F (2006) p63 protects the female germ line during meiotic arrest. Nature. 444:624-628.

Štambuk, A, Šrut, M, Šatović, Z, Tkalec, M, Klobučar, GIV (2013) Gene flow vs. pollution pressure: Genetic diversity of Mytilus galloprovincialis in eastern Adriatic. Aquatic Toxicology. 136-137:22-31.

Tanguy, A, Bierne, N, Saavedra, C, Pina, B, Bachère, E, Kube, M, Bazin, E, Bonhomme, F, Boudry, P, Boulo, V, Boutet, I, Cancela, L, Dossat, C, Favrel, P, Huvet, A, Jarque, S, Jollivet, D, Klages, S, Lapègue, S, Leite, R, Moal, J, Moraga, D, Reinhardt, R, Samain, J-F, Zouros, E, Canario, A (2008) Increasing genomic information in bivalves through new EST collections in four species: Development of new genetic markers for environmental studies and genome evolution. Gene. 408:27-36.

Thompson, JD, Gibson, TJ, Plewniak, F, Jeanmougin, F, Higgins, DG (1997) The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research. 25:4876-4882.

Tozluoğlu, M, Karaca, E, Haliloglu, T, Nussinov, R (2008) Cataloging and organizing p73 interactions in cell cycle arrest and apoptosis. Nucleic Acids Research. 36:5033-5049.

Venier, P, Pallavicini, A, De Nardi, B, Lanfranchi, G (2003) Towards a catalogue of genes transcribed in multiple tissues of Mytilus galloprovincialis. Gene. 314:29-40.

Vukmirovic, M, Bihari, N, Zahn, RK, Muller, WEG, Batel, R (1994) DNA damage in marine mussel Mytilus galloprovincialis as a biomarker of environmental contamination. Marine Ecology Progress Series. 109:165-172.

Wang, B, Xiao, Z, Ko, HL, Ren, E-C (2010) The p53 response element and transcriptional repression. Cell Cycle. 9:870-879.

Wasserman, WW, Sandelin, A (2004) Applied bioinformatics for the identification of regulatory elements. Nature reviews. Genetics. 5:276-287.

Williamson, EA, Wolf, I, O'Kelly, J, Bose, S, Tanosaki, S, Koeffler, HP (2005) BRCA1 and FOXA1 proteins coregulate the expression of the cell cycle-dependent kinase inhibitor p27Kip1. Oncogene. 25:1391-1399.

Yang, A, Kaghad, M, Caput, D, McKeon, F (2002) On the shoulders of giants: p63, p73 and the rise of p53. Trends in Genetics. 18:90-95.

Yang, A, Kaghad, M, Wang, Y, Gillett, E, Fleming, MD, Dötsch, V, Andrews, NC, Caput, D, McKeon, F (1998) p63, a p53 Homolog at 3q27–29, Encodes Multiple Products with Transactivating, Death-Inducing, and Dominant-Negative Activities. Molecular Cell. 2:305-316.

Zhou, JX, Lee, CH, Qi, CF, Wang, H, Naghashfar, Z, Abbasi, S, Morse, HC (2009) IFN Regulatory Factor 8 Regulates MDM2 in Germinal Center B Cells. Journal of immunology (Baltimore, Md. : 1950). 183:3188-3194.

8. PRILOG

Prilog 1. Izvedeni proteinski sljedovi gena porodice p53 *M. galloprovincialis*

>TAp63alfa (619 ak, 69.19 kDa, pI=6.83)

MPSHNYNPGCRFGKFTHHHPIMSQASVSTTCTPSGPPMSQETFEYLWNTLGEVTQEG
GYTNITSKESIDYAFSEAEDETSISVEKYRITSNDSISDLLNPIIGQTTTASSMSPDSQTNI
IGSSASSPYNDTITSPPPYSPHTSMQSPIPSVPSNTDYPGDYGFTISFSQPSKETKSTTWT
YSESLKKLYVRMATTCPIRFKCLRQPPQGCVIRAMPIFMKPEHVQEPVKRCPNHATSK
EHNENHPAPTHLCRCEHKLAKFVEDPYTSRQSVLIPHEIPQAGSEWVTNLFQFMCLGS
CVGGPNRRPIQIVLTLEKDNQVLGRRAVEVRICACPGRDRKADEKAALPPCKQSPKK
GQKVNIINEITTVTPGGKKRKAEDEPFTLSVRGRENYEILCRLRDSLELSSMVPQNQID
VYKQKQLDTNRQSLPTTSTARVVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQ
NGVVENHGNIKEELMANGDHSISNWLTTLGLSAYIDNFHQQNLFSMEQLDDFTVEDL
QKMRIGTSHRNKIWKALVEFHSESITISDSQSLQRDVSTASTISMTSQASISQNSQNSTY
CPGYYEVTRYTFKHTVSYTKVDERSPKRAKVD

>TAp63beta (498 ak, 55.37 kDa, pI=7.88)

MPSHNYNPGCRFGKFTHHHPIMSQASVSTTCTPSGPPMSQETFEYLWNTLGEVTQEG
GYTNITSKESIDYAFSEAEDETSISVEKYRITSNDSISDLLNPIIGQTTTASSMSPDSQTNI
IGSSASSPYNDTITSPPPYSPHTSMQSPIPSVPSNTDYPGDYGFTISFSQPSKETKSTTWT
YSESLKKLYVRMATTCPIRFKCLRQPPQGCVIRAMPIFMKPEHVQEPVKRCPNHATSK
EHNENHPAPTHLCRCEHKLAKFVEDPYTSRQSVLIPHEIPQAGSEWVTNLFQFMCLGS
CVGGPNRRPIQIVLTLEKDNQVLGRRAVEVRICACPGRDRKADEKAALPPCKQSPKK
GQKVNIINEITTVTPGGKKRKAEDEPFTLSVRGRENYEILCRLRDSLELSSMVPQNQID
VYKQKQLDTNRQSLPTTSTARVVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQ
NGVVENHGNIKEELMANGDHSISKTYKR

>TAp63gamma (455 ak, 51.14 kDa, pI=8.99)

MPSHNYNPGCRFGKFTHHHPIMSQASVSTTCTPSGPPMSQETFEYLWNTLGEVTQEG
GYTNITSKESIDYAFSEAEDETSISVEKYRITSNDSISDLLNPIIGQTTTASSMSPDSQTNI
IGSSASSPYNDTITSPPPYSPHTSMQSPIPSVPSNTDYPGDYGFTISFSQPSKETKSTTWT
YSESLKKLYVRMATTCPIRFKCLRQPPQGCVIRAMPIFMKPEHVQEPVKRCPNHATSK
EHNENHPAPTHLCRCEHKLAKFVEDPYTSRQSVLIPHEIPQAGSEWVTNLFQFMCLGS
CVGGPNRRPIQIVLTLEKDNQVLGRRAVEVRICACPGRDRKADEKAALPPCKQSPKK
GQKVNIINEITTVTPGGKKRKAEDEPFTLSVRGRENYEILCRLRDSLELSSMVPQNQID
VYKQKQLDTNRQWLSMILARENKNKLMKKVKRPQHRPGIKSRT

>deltaNp63alfa (535 ak, 60.03 kDa, pI=8.74)

MIKFERTGFTTYRLNPIIGQTTTASSMSPDSQTNIIGSSASSPYNDTITSPPPYSPHTSM
QSPIPSVPSNTDYPGDYGFTISFSQPSKETKSTTWTYSESLKKLYVRMATTCPIRFKCL
RQPPQGCVIRAMPIFMKPEHVQEPVKRCPNHATSKEHNENHPAPTHLCRCEHKLAKF
VEDPYTSRQSVLIPHEIPQAGSEWVTNLFQFMCLGSCVGGPNRRPIQIVLTLEKDNQV
LGRRAVEVRICACPGRDRKADEKAALPPCKQSPKKGQKVNIINEITTVTPGGKKRKA
EDEPFTLSVRGRENYEILCRLRDSLELSSMVPQNQIDVYKQKQLDTNRQSLPTTSTAR
VVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQNGVVENHGNIKEELMANGDH
SISNWLTTLGLSAYIDNFHQQNLFSMEQLDDFTVEDLQKMRIGTSHRNKIWKALVEF
HSESITISDSQSLQRDVSTASTISMTSQASISQNSQNSTYCPGYYEVTRYTFKHTVSYTK
VDERSPKRAKVD

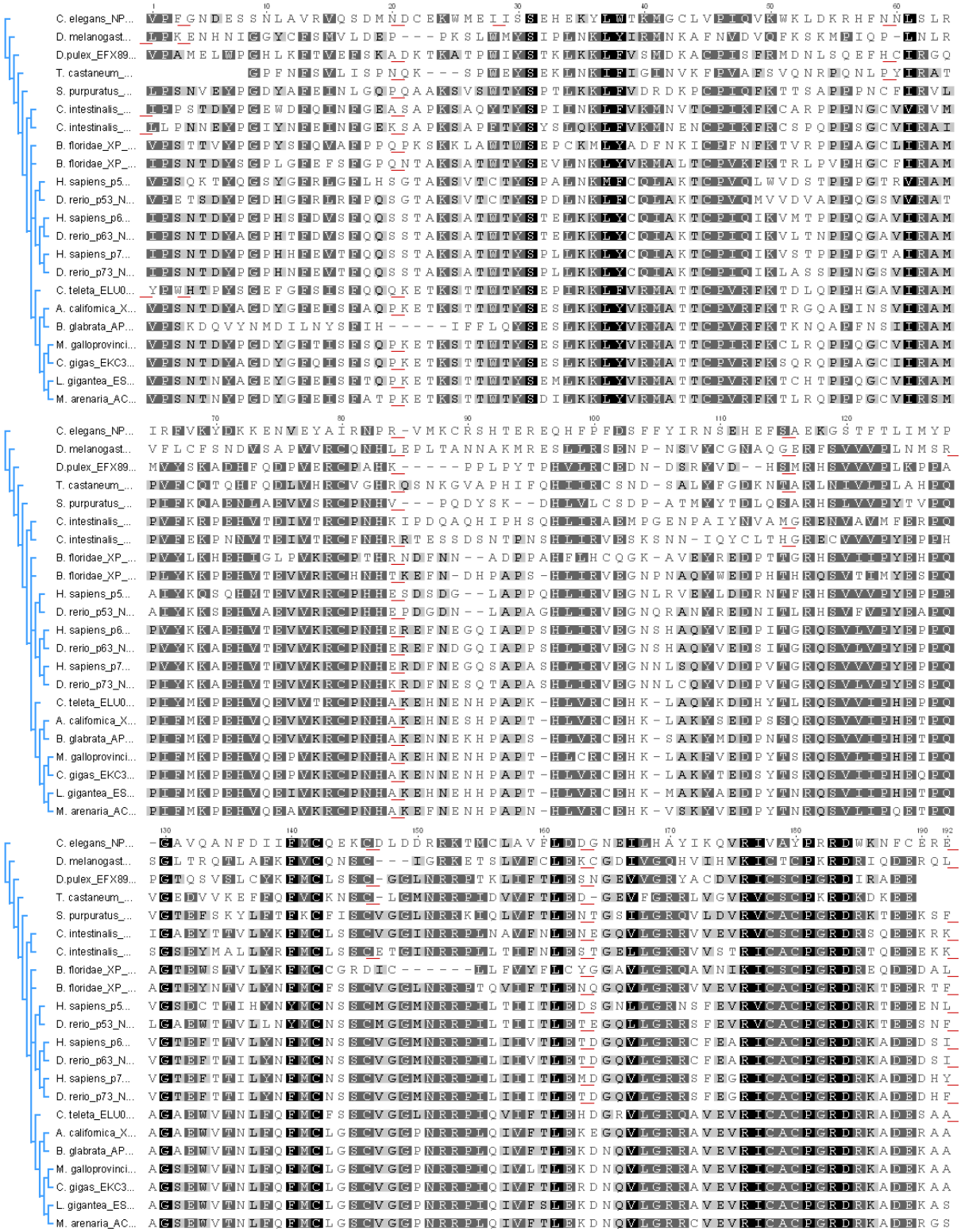
>deltaNp63beta (414 ak, 46.21 kDa, pI=9.11)

MIKFERTGFTTYRLNPIIGQTTTASSMSPDSQTNIIGSSASSPYNDTITSPPPYSPHTSM
QSPIPSVPSNTDYPGDYGFTISFSQPSKETKSTTWTYSESLKKLYVRMATTCPIRFKCL
RQPPQGCVIRAMPIFMKPEHVQEPVKRCPNHATSKEHNENHPAPTHLCRCEHKLAKF
VEDPYTSRQSVLIPHEIPQAGSEWVTNLFQFMCLGSCVGGPNRRPIQIVLTLEKDNQV
LGRRAVEVRICACPGRDRKADEKAALPPCKQSPKKGQKVNIINEITTVTPGGKKRKA
EDEPFTLSVRGRENYEILCRLRDSLELSSMVPQNQIDVYKQKQLDTNRQSLPTTSTAR
VVTLPTHDNTPITIQGEGRQTTLPFTADLNGQVTSSQNGVVENHGNIKEELMANGDH
SISKTYKR

>deltaNp63gamma (371 ak, 41.97 kDa, pI=9.61)

MIKFERTGFTTYRLNPIIGQTTTASSMSPDSQTNIIGSSASSPYNDTITSPPPYSPHTSM
QSPIPSVPSNTDYPGDYGFTISFSQPSKETKSTTWTYSESLKKLYVRMATTCPIRFKCL
RQPPQGCVIRAMPIFMKPEHVQEPVKRCPNHATSKEHNENHPAPTHLCRCEHKLAKF
VEDPYTSRQSVLIPHEIPQAGSEWVTNLFQFMCLGSCVGGPNRRPIQIVLTLEKDNQV
LGRRAVEVRICACPGRDRKADEKAALPPCKQSPKKGQKVNIINEITTVTPGGKKRKA
EDEPFTLSVRGRENYEILCRLRDSLELSSMVPQNQIDVYKQKQLDTNRQWLSMILARE
NKNKLMKKVKRPQHRPGIKSRT

Prilog 2. Sravnjenje središnjih očuvanih regija proteinskih sljedova (DBD) korištenih za filogenetsku analizu



9. ŽIVOTOPIS

Ana Baričević je rođena 5. travnja 1982. u Dubrovniku gdje je završila osnovnu i srednju školu. Prirodoslovno-matematički fakultet Sveučilišta u Zagrebu, studij biologije, smjer molekularna biologija, upisala je 2000. godine. Za vrijeme trajanja studija prima državnu stipendiju Ministarstva znanosti, obrazovanja i sporta. Diplomski rad naslova „Apoptoza embrionalnih stanica miša uzrokovana UV-svjetlosti“ izradila je na Zavodu za molekularnu biologiju Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu i obranila 2007.godine. Iste godine zapošljava se na Institutu "Ruđer Bošković", Centar za istraživanje mora u Rovinju, gdje i danas radi kao znanstvena novakinja u Laboratoriju za morsku molekularnu biologiju. Područje interesa joj je genetika i ekologija morskih organizama.

Koautor je na 4 znanstvena rada i sudjelovala je na više znanstvenih skupova i radionica.

Znanstveni radovi:

Blažina, M., Najdek, M., Fuks, D., Ruso, A., Štifanić, M., Pavlinić, D. (2009). Characterisation and differentiation of oligotrophic waters by culturable particle-attached and free-living bacterial communities. *Ecological Indicators*, *9*(6), 1265-1270.

Štifanić, M., Mičić, M., Ramšak, A., Blašković, S., Ruso, A., Zahn, R. K., Batel, R. (2009). p63 in *Mytilus galloprovincialis* and p53 family members in the phylum Mollusca. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *154*(3), 264-273.

Blažina, M., Najdek, M., Ruso, A., Fuks, D. (2010). Adaptational changes in cellular fatty acids of cultured bacteria as a response to trophic differences. *Acta Adriatica*, *51*(2), 131-140.

Baričević, A., Štifanić, M., Hamer, B., Batel, R. (2015). p63 gene structure in the phylum Mollusca. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *186*, 51-58.

Sudjelovanje na skupovima:

Baričević, Ana; Štifanić, Mauro; Treursić, Barbara; Hamer, Bojan; Batel, Renato
Genomic structure of p63, the ancestral gene of the p53 family, in mussel *Mytilus galloprovincialis* // *3rd Congress of Croatian Geneticists*
Krk, Hrvatska, 2012 (usmeno pripćenje, prvi autor)

Baričević, Ana
p63 gene in mussel *Mytilus galloprovincialis*-potential biomarker of genotoxic stress in environmental studies // *Summer School*
Rovinj, Hrvatska, 2014 (usmeno pripćenje, prvi autor)

Štifanić, Mauro; Baričević, Ana; Batel Renato
Non-vertebrate animals have only one gene of the p53 family which is most similar to vertebrate p63 // *27th ESCPBnew Congress*
Alessandria, Italija, 2010 (usmeno pripćenje, nije prvi autor)

Štifanić, Mauro; Mičić, Milena; Ramšak, Andreja; Blašković, Sanja; Ruso, Ana; Batel, Renato
p63 in *Mytilus galloprovincialis* and p53 family members in the phylum mollusca // *10th Croatian Biological Congress*
Osijek, Hrvatska, 2009 (usmeno pripćenje, nije prvi autor)

Baričević, Ana; Perusco, S.Victor; Medić,Nikola; Hamer Bojan
Population analyses (cox1 gene) and bile PAH metabolites content of commercial fishes along Adriatic coast // *4th Conference on Conservation Physiology of Marine Fishes*
Thessaloniki, Grčka, 2014 (poster)

Korlević, Marino; Durmiši, Emina; Treursić, Barbara; Baričević, Ana; Hamer, Bojan
Mytilus galloprovincialis ferritin: gene and cDNA sequence analysis // *46th European Marine Biology Symposium*
Rovinj, Hrvatska, 2011 (poster)

Ana Baričević
Marine Evolutionary Genomics and Proteomics
Vigo, Španjolska, 2014 (radionica)

Ana Baričević
International School of Conservation Biology (iSCB)
Rovinj, Hrvatska, 2011 (radionica)

Ana Baričević
Writing effectively about your research
Zagreb, Hrvatska, 2010 (radionica)

Ana Baričević
BIOMINTEC-Biomineralization:Understanding of basic mechanisms for the design of novel strategies in nanobiotechnology
Rovinj, Hrvatska, 2009 (radionica)