

Sveučilište Josipa Jurja Strossmayera u Osijeku
Sveučilište u Dubrovniku
Institut Ruđer Bošković
Sveučilišni poslijediplomski interdisciplinarni doktorski studij
Molekularne bioznanosti

Eva Šatović

**Organizacija ponovljenih sekvenci DNA u genomima
tri vrste školjkaša**

doktorski rad

Zagreb, 2013.

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište Josipa Jurja Strossmayera u Osijeku
Sveučilište u Dubrovniku
Institut Ruđer Bošković
Sveučilišni poslijediplomski interdisciplinarni
doktorski studij Molekularne bioznanosti

Doktorski rad

Znanstveno područje: Prirodne znanosti
Znanstveno polje: Biologija

Organizacija ponovljenih sekvenci DNA u genomima tri vrste školjkaša

Eva Šatović

Rad je izrađen u: Laboratoriju za strukturu i funkciju heterokromatina,
Institut Ruđer Bošković, Zagreb

Mentor/i: dr. sc. Miroslav Plohl

Kratki sažetak doktorskog rada:

Značajan dio eukariotskog genoma čine uzastopno i raspršeno ponovljene sekvence DNA. One su predmetom istraživanja ekološki i komercijalno značajnih školjkaša: *Donax trunculus*, *Ruditapes decussatus* i *Ruditapes philippinarum*. Rezultati donose informacije o nizu novih mobilnih elemenata i satelitnih DNA, kao i o njihovoj međusobnoj povezanosti u jedinstvenu kompleksnu mrežu nekodirajućih sekvenci.

Broj stranica: 139

Broj slika: 47

Broj tablica: 7

Broj literaturnih navoda: 128

Jezik izvornika: hrvatski

Ključne riječi: satelitna DNA, mobilni elementi, školjkaši

Datum obrane: 20. svibnja 2013.

Stručno povjerenstvo za obranu:

- 1.
- 2.
- 3.
- 4.
- 5.
6. (zamjena)

Rad je pohranjen u: Gradskoj i sveučilišnoj knjižnici Osijek, Europska avenija 24, Osijek;
Sveučilištu Josipa Jurja Strossmayera u Osijeku, Trg sv. Trojstva 3, Osijek

BASIC DOCUMENTATION CARD

Josip Juraj Strossmayer University of Osijek
University of Dubrovnik
Ruđer Bošković Institute
University Postgraduate Interdisciplinary Doctoral Study of
Molecular biosciences

PhD thesis

Scientific Area: Natural sciences
Scientific Field: Biology

Organization of repetitive DNA sequences in genomes of three bivalve species

Eva Šatović

Thesis performed at: Laboratory for structure and function of heterochromatin,
Ruđer Bošković Institute, Zagreb

Supervisor/s: dr. sc. Miroslav Plohl

Short abstract:

Significant part of eukaryotic genomes constitute tandem and interspersed repetitive DNA sequences. They are explored in bivalve species of ecological and commercial importance: *Donax trunculus*, *Ruditapes decussatus* and *Ruditapes philippinarum*. Results bring information about new mobile elements and satellite DNAs, and about their integration into a complex network of noncoding sequences.

Number of pages: 139

Number of figures: 47

Number of tables: 7

Number of references: 128

Original in: croatian

Key words: satellite DNA, mobile elements, bivalves

Date of the thesis defense: 20. May 2013.

Reviewers:

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.(substitute)

Thesis deposited in: City and University Library of Osijek, Europska avenija 24, Osijek;
Josip Juraj Strossmayer University of Osijek, Trg sv. Trojstva 3, Osijek

Ovaj doktorski rad je izrađen u Laboratoriju za strukturu i funkciju heterokromatina Instituta Ruđer Bošković u Zagrebu, pod vodstvom dr.sc. Miroslava Plohla, u sklopu Sveučilišnog poslijediplomskog interdisciplinarnog doktorskog studija Molekularne bioznanosti.

Sadržaj

1. Uvod.....	1
1.1. Organizacija ponovljenih sekvenci DNA u genomima eukariotskih organizama.....	1
1.2. Evolucijska dinamika satelitnih DNA.....	4
1.3. Raspršeno ponovljene sekvence.....	6
1.4. Evolucijska važnost mobilnih elemenata.....	10
1.5. Povezanost satelitnih sekvenci i mobilnih elemenata.....	12
1.6. Osnovna obilježja školjkaša i njihovog genoma.....	14
1.7. Obilježja i evolucija ponovljenih sekvenci DNA kod školjkaša.....	17
1.8. Cilj.....	24
2. Materijali i metode.....	25
2.1. Izolacija genomske DNA školjkaša.....	25
2.2. Razgradnja genomske DNA restrikcijским endonukleazama.....	25
2.3. Umnažanje ciljnih nukleotidnih sljedova uz korištenje lančane reakcije polimerazom.....	26
2.3.1. Umnažanje Dt-BIV160 skupine sekvenci.....	26
2.3.2. Dobivanje cjelovite sekvence dugih ukloniranih genomskih fragmenata.....	28
2.3.3. Umnažanje centralnih ponavljanja DTC84.....	29
2.3.4. Detekcija i karakterizacija dodatnih elemenata DTC84.....	30
2.4. Pročišćavanje PCR produkata.....	31
2.5. Elektroforetsko odjeljivanje fragmenata DNA.....	31
2.6. Izolacija fragmenata DNA iz agaroznog gela.....	32
2.7. Ligacija fragmenata DNA u plazmidne vektore.....	32
2.8. Transformacija bakterije <i>Escherichia coli</i>	32
2.9. Uzgoj bakterijskih stanica.....	32
2.10. Izolacija plazmidne DNA.....	33
2.11. Određivanje primarne strukture DNA.....	33
2.12. Obilježavanje hibridizacijskih sonda.....	33
2.13. Hibridizacijska analiza prema Southernu.....	34
2.14. Kapljična analiza DNA (eng. <i>dot blot</i>).....	35
2.15. Analiza otisaka kolonija (eng. <i>colony lift</i>).....	35
2.16. Računalne analize nukleotidnih sekvenci.....	36

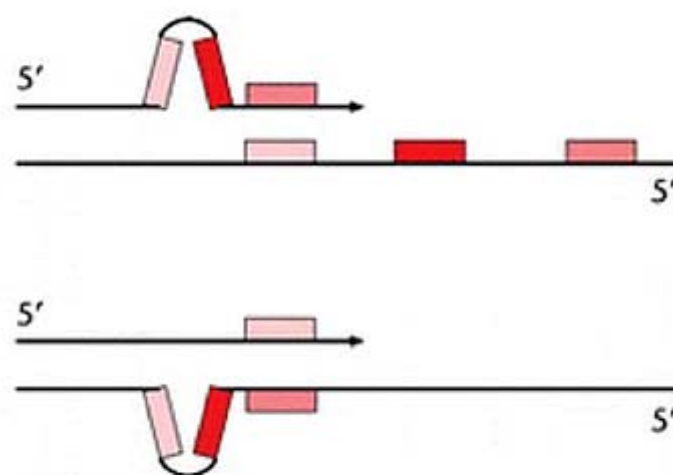
3. Rezultati.....	37
3.1. Pregled repetitivnih sekvenci u genomu školjkaša <i>Donax trunculus</i>	37
3.1.1. BIV160 porodica satelitnih DNA u školjkašu <i>Donax trunculus</i>	48
3.1.2. DTC84 element.....	53
3.1.2.1. Organizacija i udio DTC84 ponavljanja u genomu školjkaša <i>Donax trunculus</i>	63
3.1.3. Ostale sličnosti s elementima porodice pearl.....	65
3.1.4. Sličnost satelitnih DNA školjkaša <i>Donax trunculus</i> i različitih skupina mobilnih elemenata.....	67
3.1.5. Grupiranja sekvenci.....	68
3.1.6. DTC M1.....	72
3.1.7. DTC M2.....	74
3.1.8. DTC T1.....	76
3.2. Pregled repetitivnih sekvenci u genomu školjkaša <i>Ruditapes decussatus</i>	79
3.2.1. Satelitne DNA izolirane iz vrste <i>Ruditapes decussatus</i>	84
3.2.2. Element RUDI 1.....	87
3.2.3. Element D10.....	88
3.3. Pregled repetitivnih sekvenci u genomu školjkaša <i>Ruditapes philippinarum</i>	89
3.3.1. Satelitne DNA izolirane iz vrste <i>Ruditapes philippinarum</i>	97
3.3.2. Element RPHI T1.....	100
3.3.3. Element RUDI 2.....	103
3.3.4. Klon P18.....	104
3.3.5. Usporedbe repetitivnog dijela genoma školjkaša <i>R. decussatus</i> i <i>R. philippinarum</i>	106
4. Rasprava.....	107
5. Zaključci.....	121
6. Literatura.....	123
7. Sažetak.....	133
8. Summary.....	135
9. Popis kratica.....	137
10. Životopis.....	139

1. Uvod

1.1. Organizacija ponovljenih sekvenci DNA u genomima eukariotskih organizama

Genomi eukariotskih organizama sadrže visok udio nekodirajućih sekvenci DNA. Udio nekodirajuće DNA može donekle varirati među vrstama, a kod čovjeka iznosi i do 99%. Veći dio nekodirajućih sekvenci čine različiti tipovi ponovljenih sekvenci DNA, a one se mogu naći organizirane u uzastopno ili raspršeno ponovljenom obliku. Uzastopna ponavljanja čine satelitne, minisatellitne, mikrosatellitne i telomerne sekvence DNA a razlikuju se po duljini jedinice ponavljanja, mehanizmima nastanka, duljini nizova koje tvore, organizaciji i položaju na kromosomima. Raspršeno ponovljene sekvence nastaju kao posljedica premještanja mobilnih elemenata po genomu.

Mikrosatellitne sekvence DNA sastoje se od jedinica ponavljanja duljine 2 - 5 pb i tvore nizove duljine do 500 pb. Do skraćivanja ili produljivanja niza dolazi uslijed proklizavanja tijekom replikacije DNA (Slika 1.1). Duljine nizova na pojedinim lokusima mogu biti izrazito varijabilne među jedinkama, što ih čini dobrim genetičkim markerima. Mikrosatelitni nizovi se mogu naći u kodirajućim i nekodirajućim dijelovima gena pri čemu mogu utjecati na gensku ekspresiju, a njihova prisutnost je česta i blisko povezana i uz raspršeno ponovljene mobilne genetičke elemente (Ramsay i sur. 1999, Wilder i Hollocher 2000, Coates i sur. 2011).

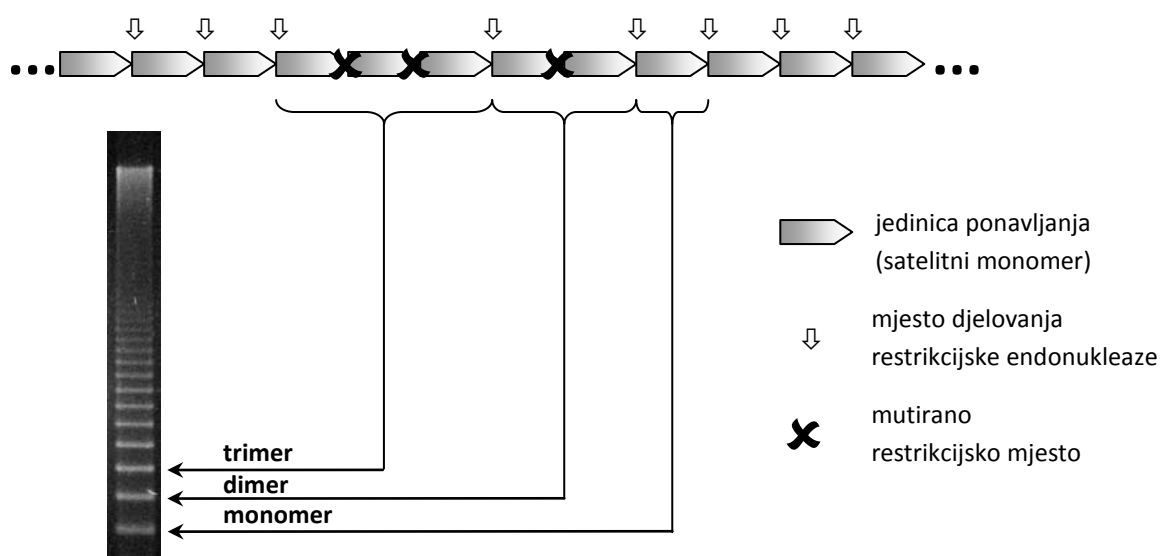


Slika 1.1: Produljenje niza nastalo pogrešnim sparivanjem uslijed stvaranja omče u vodećem lancu. (a) Skraćivanje niza nastalo pogrešnim sparivanjem uslijed stvaranja omče u lancu koji služi kao kalup. (b) (Preuzeto iz Moran i Arbor 2005)

Jedinice ponavljanja sekvenci minisatelitnih DNA iznose 10 - 100 pb i tvore nizove duljine 0,5 - 30 kilobaza. Mogu se naći na različitim kromosomskim lokacijama, pretežito u eukromatinskim područjima.

Za razliku od ove dvije skupine koje su u genomu ponovljene u umjerenom broju kopija, satelitne DNA imaju jedinice ponavljanja koje su uzastopno ponovljene u velikom broju, tvoreći u nekim slučajevima nizove i do 100 megabaza duljine. Sami monomeri mogu biti različite duljine, od nekoliko pb pa sve do 40 kb (Benedum i sur. 1986) no najčešće se kreću u rasponu od 100 do 400 pb. Zapaženo je kako se duljina monomera većine centromernih satelita kreće u rasponu od približno 150 do 210 pb i odgovara duljini nukleosomskih jedinica (Henikoff i sur. 2001). Udio pojedinih satelitskih DNA u genomu također varira u širokom rasponu od onih <1%, preko 50% (kornjaš *Tenebrio molitor*), pa čak i do 90% kod nekih vodozemaca, riba i cvjetnica (Bachmann i sur. 1991). Satelitna DNA je dobila naziv prema tome što je taj tip sekvenci bio prvotno izoliran iz dodatne, "satelitne" pruge, koja se nakon centrifugiranja u gradijentu gustoće cezij klorida izdvojila od ostatka genomske DNA zbog drugačijeg AT sastava (John 1988). Danas, nakon što su opisane brojne satelitne DNA, i dalje stoji tvrdnja da se one često razlikuju svojim AT sastavom od prosječnog AT sastava genoma. Većina opisanih satelitnih DNA je AT-bogata, no mogu se naći i GC-bogate iznimke (npr. Petrović i Plohl 2005).

Osnovna organizacija satelitnih DNA u dugačke nizove monomera uzastopno ponovljenih u orijentaciji glava-rep omogućila je i lako utvrđivanje tih sekvenci u genomima nakon cijepanja genomske DNA restrikcijskim enzimima. Nakon elektroforetskog odvajanja fragmenata DNA vidljive su karakteristične ljestvice koji odgovaraju monomeru i njegovim višekratnicima (pregled u Bachmann i sur. 1991). Naime, uzastopni monomeri obično se neznatno razlikuju u nukleotidnoj sekvenci (zbog pojave točkastih mutacija), čime može biti promijenjeno mjesto prepoznavanja restrikcijskog enzima te on na tom mjestu ne može rezati DNA. U takvom će se slučaju na elektroforetskom gelu moći primijetiti ljestvica sastavljena od monomera, dimera i ostalih multimera (Slika 1.2). Prisutnost većeg broja pruga u sklopu ljestvice može se uzeti i kao indikacija da su razlike među monomerima učestalije, dok ljestvica koja se sastoji isključivo od monomera i dimera ukazuje na nizove sastavljene od ponavljanja vrlo visoke homologije.



Slika 1.2: Osnovna organizacija jedinica ponavljanja satelitne DNA te princip pojavljivanja karakteristične ljestvice nakon razgradnje genomske DNA restriktivnim enzimima (preuzeto iz Mravinac 2004).

Ponekad susjedni monomeri (ili više monomera u nizu) formiraju novu jedinicu ponavljanja, koja se zove jedinicom organizacije višeg reda (eng. *higher-order repeat*, skraćeno HOR). Glavna odlika HOR-a jest jasno razlučivanje njezinih sastavnih elemenata te veća sličnost između HOR-a nego što je sličnost između susjednih monomera unutar pojedinog HOR-a. Najbolje opisana organizacija višeg reda jest ona karakteristična za α -satelitnu DNA primata. Temelji se na jedinicama ponavljanja duljine 171 pb koje formiraju jedinice HOR-a raznih duljina, od dimernih pa sve do 16-mernih struktura (Warburton i sur. 1993).

Na kromosomima, satelitne DNA se najčešće nalaze u područjima (peri)centromernog ili subteloernog heterokromatina a iznimno u eukromatinskim područjima (Martinez-Lage i sur. 1994). Satelitne DNA smještene u centromernim i pericentromernim heterokromatinskim područjima dovedene su u vezu s jednom od osnovnih bioloških funkcija, pravilnim razdvajanjem kromosoma u mejozi i mitozu (Schueler i sur. 2001, Sun i sur. 2003). Predloženo je da promjene u centromernoj DNA preko promjena u centromernim proteinima mogu dovesti do specijacije, tj. procesa odvajanja vrsta (Meštrović i sur. 1998, Henikoff i sur. 2001). Dobro poznati DNA motiv uključen u interakciju s centromernim proteinima je 17 pb

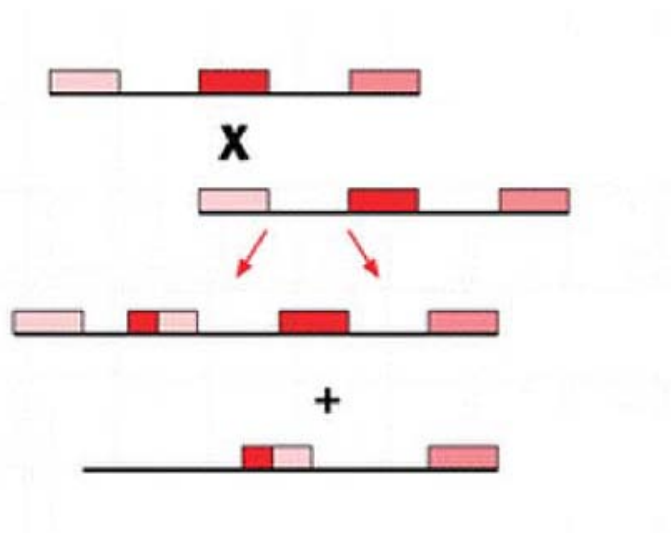
dug slijed CENP-B, nađen unutar dijela varijanti monomera α -satelitne DNA čovjeka. Taj motiv prepoznaje protein CENP-B (Masumoto i sur. 1989) i vjerojatno sudjeluje pri formiranju centromernog heterokromatina (Ohzeki i sur. 2002). Kod sisavaca su motiv i protein CENP-B visoko konzervirani. Mišji protein CENP-B tako pokazuje 92% sličnosti s ljudskim te se također veže na motiv koji nalikuje motivu CENP-B u području centromerne satelitne DNA (Sullivan and Glass 1991). Varijante tog motiva nađene su i kod beskralješnjaka, uključujući i školjkaše (Canapa i sur. 2000). Uočena je značajna sličnost između proteina CENP-B i transpozaze elementa *pogo* u vrsti *Drosophila melanogaster* (Tudor i sur. 1992), što bi ukazivalo na njegovo porijeklo od superfamilije Tc1/mariner transpozaza.

Satelitne DNA su tradicionalno smatrane potpuno transkripcijski neaktivnima no transkripti satelitnih DNA pronađeni su u određenom broju organizama. Pokazano je i kako kratki transkripti nekih satelitnih DNA upravljaju procesima izgradnje heterokromatina (Volpe i sur. 2002, Verdel i sur. 2004) te postoje indikacije kako satelitne DNA mogu igrati regulatornu ulogu u eukariotskim genomima (pregled u Ugarković 2005).

1.2. Evolucijska dinamika satelitnih DNA

U eukariotskim genomima istovremeno postoji veći broj različitih satelitnih DNA koje se međusobno razlikuju po nukleotidnom slijedu monomera i udjelu u genomu, te zajedno čine vrsno specifični satelitni profil (Ugarković i Plohl 2002, Plohl i sur. 2008). Još 1977. godine Fry i Salser predložili su "teoriju biblioteke" po kojoj srodne vrste dijele čitav skup satelitnih sekvenci, prisutan još u zajedničkom pretku. Bilo koji član te biblioteke može se unutar pojedine vrste umnožiti i na taj način definirati vrsno specifični profil. Prvi eksperimentalni podaci koji u potpunosti podržavaju tu pretpostavku dobiveni su proučavanjem satelitnih DNA u rodu *Palorus* (Meštrović i sur. 1998), gdje je svaka od proučavane četiri vrste sadržavala iste satelitne porodice ali prisutne u različitom broju kopija. Danas se nastanak pojedinih satelitnih profila redovito razmatra u svjetlu teorije biblioteke.

Premda stupanj sličnosti satelitnih DNA između srodnih vrsta može znatno varirati, evolucija ovih sekvenci slijedi načela “usklađene evolucije” (Dover 1986). Usklađena evolucija ponovljenih sekvenci rezultira većom srodnošću satelitnih monomera unutar vrste, nego što je sličnost između monomera porijeklom iz različitih vrsta. Takvo stanje posljedica je homogenizacije novonastalih varijanti monomera unutar genoma jedinke, kao i istovremenog fiksiranja novonastale varijante među jedinkama reproduktivno povezane zajednice (Dover, 1986). Među mehanizmima homogenizacije svakako su najvažniji nejednoliko prekriženje (eng. *crossing-over*) i genska konverzija, no također je značajno i umnažanje "kotrljajućim krugom" te mehanizmi zasnovani na transpoziciji (Plohl i sur. 2012). Genska konverzija je proces kojim se mutacije nastale u jednom monomeru šire na drugi, ne mijenjajući time broj ponovljenih jedinica. Nejednoliko prekriženje (Slika 1.3) je međutim recipročni rekombinacijski proces koji širi promjene mijenjajući pritom i broj kopija te se smatra odgovornim za česte promjene u duljini ponavljajućih nizova (Smith 1976). Budući da su otkrivene i izvankromosomske kružne molekule DNA sastavljene od satelitnih jedinica (Cohen i sur. 2003), pretpostavlja se da bi se dugi satelitni nizovi mogli umnažati vankromosomski, putem replikacije "kotrljajućim krugom", iza čega bi uslijedila njihova reinsercija u genom (Ohki i sur. 1995).



Slika 1.3: Nejednoliko „*crossing-over*“ koji dovodi do produljenja jednog niza ponavljanja i skraćivanja drugog. (Preuzeto iz Moran i Arbor 2005)

Struktura nekih satelitnih jedinica ponavljanja koja podsjeća na strukturnu organizaciju transponirajućih elemenata, sugerira transpoziciju kao mogući amplifikacijski mehanizam (pregled u Plohl i sur. 2012).

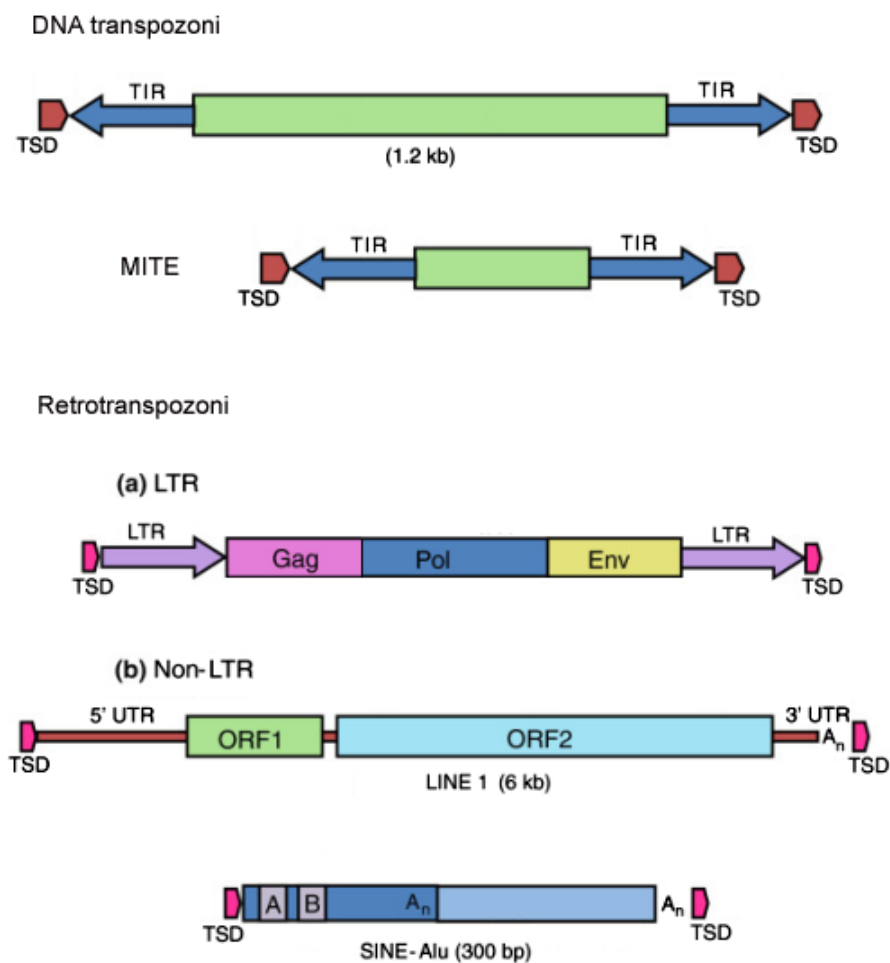
1.3. Raspršeno ponovljene sekvence

Raspršeno ponovljene sekvence nastaju kao posljedica premještanja mobilnih elemenata po genomu. Mobilne ili transponirajuće elemente prve je uočila Barbara McClintock, i to nekoliko godina prije nego je uopće razriješena struktura same molekule DNA (McClintock 1950). Kasnije su transponirajući elementi pronađeni i kod prokariotskih i kod eukariotskih organizama a mogu se svrstati u dvije glavne porodice: retrotranspozoni (elementi klase I) i DNA transpozoni (elementi klase II) (Slika 1.4.). Kod obje skupine možemo razlikovati elemente koji su autonomni i one koji to nisu, to jest one koji za premještanje po genomu koriste enzime autonomnih elemenata. Važno je naglasiti da potpuni, autonomni elementi, čine manjinski udio kod većine genoma koje nastanjuju, dok prevladavaju neautonomni elementi ili samo fragmenti izvornih, cjelovitih kopija.

Uzevši u obzir njihovu strukturu i mehanizme transpozicije elementi klase I se mogu dalje podijeliti na LTR retrotranspozone i non-LTR retrotranspozone. „LTR“ skraćenica dolazi od eng. „*long terminal repeats*“ i označava dugačke sekvence koje se nalaze na jednom i drugom kraju elementa ponovljene u direktnoj orijentaciji, a mogu biti duljine od ~100 pb do 5 kb. Ovi elementi se premještaju reverznom transkripcijom nakon koje slijedi ugradnja uz pomoć vlastite integraze. Kod LTR retrotranspozona se između rubnih ponavljanja mogu naći barem 2 otvorena okvira čitanja (eng. *open reading frame*, skraćeno ORF), *gag* i *pol*, nalik onima prisutnim kod retrovirusa, koje također možemo ubrojiti u posebnu skupinu retrotranspozona. Lokus *gag* kodira proteine uključene u formiranje kapside, a *pol* kodira 4 proteina bitna za kopiranje i integraciju elementa (proteazu, reverznu transkriptazu, RNAzu H i integrazu). Za razliku od retrovirusa, treći, *env* gen, je kod LTR retrotranspozona ili nefunkcionalan ili nedostaje. Daljnja subklasifikacija LTR retrotranspozona uključuje dvije osnovne grupe, Ty1-copia i Ty3-gypsy (Capy 2005).

Važno je spomenuti da transkripcija LTR retrotranspozona ne korelira nužno s njihovom insercijom. Uzevši u obzir da replikacijski ciklus LTR retrotranspozona uključuje 4 koraka: transkripciju, translaciju, reverznu transkripciju i integraciju – regulacija bilo kojeg od tih koraka može ograničiti transpoziciju. Na primjer, unatoč tome što nastaje velik broj transkripata kvašćevog retrotranspozona Ty1, nove su insercije iznimno rijetke. Naime, uslijed inhibicije, na svakih 14000 transkripata elementa Ty1 nastaje samo jedna cDNA (Curcio i Garfinkel 1999). Postoje i suprotni primjeri gdje je širenje retrotranspozona takvo da može imati značajan utjecaj na genom i njegovu veličinu. SanMiguel i sur. (1998) su pokazali da su nagle eksplozije LTR retrotranspozona nakupljenih u intergenskim regijama tijekom 6 milijuna godina postepeno udvostučile veličinu genoma kukuruza.

U sklopu porodice LTR retrotranspozona postoji i jedna posebna grupa elemenata pod nazivom TRIM (eng. *terminal-repeat retrotransposons in miniature*), koji su izvorno nađeni kod biljaka (Arabidopsis Genome Initiative 2000). Radi se o manjim elementima bez mogućnosti kodiranja proteina vezanih uz mobilnost. Elementi TRIM sadrže TDR (eng. *terminal direct repeat*) sljedove duljine 100 - 250 pb, između kojih se nalazi unutarnja domena duljine 100 - 300 pb. Mogu se naći na različitim lokacijama u genomu, u velikom broju kopija i s visoko očuvanom homologijom među TDR sljedovima, koja se i kod udaljenih vrsta zadržava na razini oko 60 - 75%. Unutarnja domena započinje s PBS (eng. *primer binding site*) slijedom, a završava polipurinskim traktom. Često se na krajevima elementa mogu naći očuvane duplikacije na mjestu insercije (eng. *target site duplication*, skraćeno TSD) duljine 5 pb, kao znak njihove mobilnosti. S obzirom na njihovu prisutnost u promotorskim i intronskim sekvencama, kao i u kodirajućim sljedovima te EST (eng. *expressed sequence tag*) bazama pojedinih biljnih vrsta čini se da su TRIM elementi aktivno uključeni u restrukturiranje biljnih genoma (Witte i sur. 2001). Također, pokazali su se kao ciljna mjesta za daljnje retrotranspozicijske insercije. Osim u biljnom svijetu TRIM elementi su zasad nađeni i u životinjskom svijetu, npr. kod mrava *Pogonomyrmex barbatus* (Zhou i Cahan 2012).



Slika 1.4: Osnovna podjela raspršeno ponovljenih sekvenci

Druga skupina elemenata klase I su non-LTR retrotranspozoni u koje ubrajamo LINE (eng. *Long Interspersed Nuclear Element*) sekvence. Na 5' i 3' krajevima ovih elemenata nalaze se UTR (eng. *untranslated region*), netranslatirane regije, pri čemu 5'-UTR sadrži promotorsku sekvencu, a 3'-UTR završava s poli-A slijedom. Ovi elementi sadrže promotor za RNA polimerazu II i 2 otvorena okvira čitanja, pri čemu proteinski produkti prvog imaju RNA-vezujuću sposobnost a produkti drugog pokazuju aktivnost endonukleaze i reverzno-transkriptaze. Puna duljina LINE elemenata je oko 6 kb, međutim, reverzna transkripcija često ne dosegne do 5' kraja što rezultira mnogim skraćenim i inaktivnim varijantama. Kod čovjeka

prevladavaju LINE elementi porodice L1, a prisutni su i u humanoj centromernoj sekvenci. Ondje su visoko homogenizirani nizovi α -satelitne DNA povremeno prekinuti s „mlađim“ elementima L1, dok se nastavno na njih nalaze divergentnije varijante α -satelitne DNA sa „starijim“ i krnjim varijantama elemenata L1 (Schueler i sur. 2001). Mehanizmi transpozicije elemenata L1 uključuju iste procese spomenute kod LTR retrotranspozona.

SINE (*Short Interspersed Nuclear Element*) su relativno kratke (<700 pb) neautonomne retrosekvence koje se prepisuju sa vlastitog promotora uz pomoć stanične RNA polimeraze III, dok njihova reverzna transkripcija ovisi o enzimima autonomnih LINE sekvenci. 5' kraj SINE elementa je nastao prepisivanjem neke od staničnih RNA: tRNA, 7SL RNA ili 5S rRNA. U tom dijelu se nalazi promotor za RNA polimerazu III u čijem sklopu su prisutni „box A“ i „box B“ motivi. Drugi dio segmenta može biti porijeklom od LINE sekvenci. Između ta dva građevna bloka ponekad može biti prisutna i središnja (*core*) domena, koja obično ostaje očuvana i kod međusobno udaljenih SINE porodica. Kao i LINE, SINE završava s poli-A slijedom. SINE sekvence prisutne u humanom genomu s više od 1 000 000 kopija su Alu sljedovi. Oni su porijeklom od 7SL RNA, koja je sastavna komponenta SRP (eng. *signal recognition particle*) kompleksa u stanici (Maraia i Sarrowa, 1995), dok su MIR sljedovi (eng. *mammalian interspersed elements*) porijeklom od tRNA. SINE elementi se transponiraju istim mehanizmom kao L1 elementi, pri čemu im je za aktivnost potreban produkt ORF2 (Belancio i sur. 2008). Dok LINE sekvence pokazuju snažnu naklonost AT-bogatim regijama i regijama bez prisutnosti gena, SINE se češće mogu naći u GC-bogatim regijama i genskim regijama (genske su obično i GC bogate).

Klasa II raspršeno ponovljenih elemenata obuhvaća DNA transpozone. Oni na svojim krajevima sadrže sekvence ponovljene u obrnutoj orijentaciji, TIR (eng. *terminal inverted repeat*) sekvence. DNA transpozoni kodiraju enzim transpozazu koja se veže za invertirane ponovljene sekvence (ili pored njih) i omogućuje transpoziciju jednim od sljedećih mehanizama: nereplikativnom, konzervativnom ili replikativnom transpozicijom (Kleckner 1990). Kod nereplikativne i konzervativne transpozicije je mobilni element izrezan sa svoje lokacije na kromosomu i ubačen na novu lokaciju (tzv. „*cut and paste*“ mehanizam) i tim procesom ne nastaju nove kopije transpozona. Nereplikativna transpozicija ipak uključuje nešto malo sinteze DNA na novom mjestu ugradnje, dok konzervativna uključuje precizne mehanizme cijepanja i ligacije, nakon kojih sinteza nije potrebna (Kleckner 1990). Kod replikativne transpozicije (tzv. „*copy and paste*“ mehanizam) transpozona se duplicira tijekom

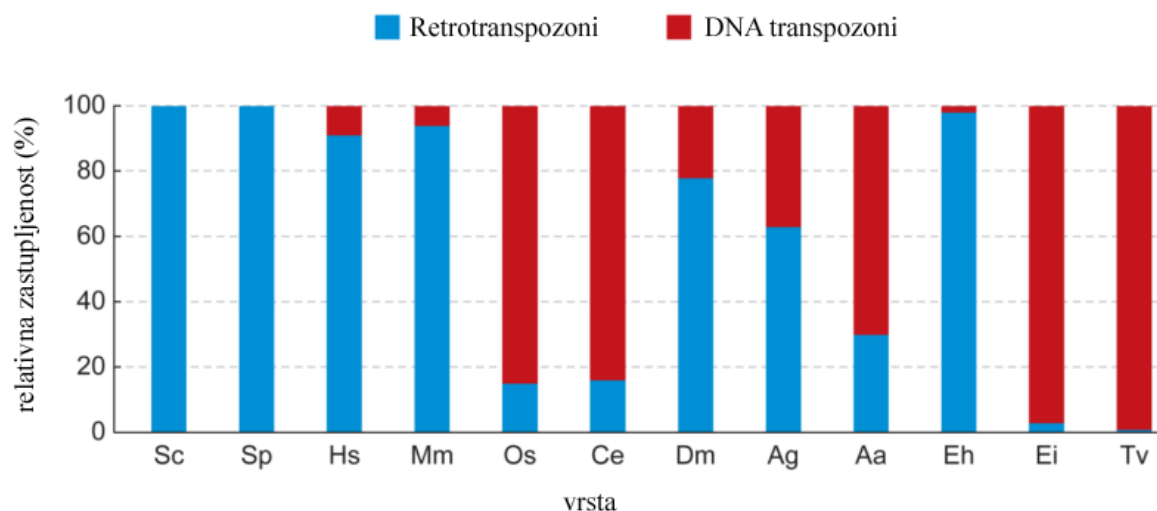
replikacije te nastaje intermedijer kod kojeg su donorsko i ciljno mjesto međusobno spojeni krajevima transpozona. Nakon razrješenja strukture, oba mjesta sadrže po jednu kopiju transpozona (Kleckner 1990). Osim ovih „klasičnih“ mehanizama (prisutni npr. kod P-elemenata), postoji još i mehanizam replikacije „kotrļajućim krugom“ kao kod Helitron elemenata (Pritham i Feschotte 2007) te mehanizmi kod kojih elementi vjerojatno koriste vlastitu DNA polimerazu, ali čiji mehanizmi transpozicije još nisu potpuno jasni; npr. Maverick elementi (Pritham i sur. 2007). Uzevši u obzir da su invertne sekvence na krajevima elementa i njihovo prepoznavanje od strane transpozaze jedini uvjeti za transpoziciju, primjerci s unutarnjim delecijama ili rearanžmanima mogu se i dalje transponirati uz pomoć transpozaze autonomnih elemenata. Predvidivo je da će na taj način autonomne kopije biti brojčano nadjačane od strane neautonomnih.

Genomi koji sadrže najveću količinu i najrazličitije vrste DNA transpozona (riža, nematode, čovjek) su ujedno i oni s najvećim brojem MITE (eng. *miniature inverted-repeat transposable element*) elemenata (Feschotte i sur. 2002, Jiang i sur. 2004). MITE su kratki (100 – 600 pb) elementi, prisutni u visokom broju kopija, očuvanih TIR sekvenci, ali bez mogućnosti autonomne transpozicije. Strukturna homogenost unutar MITE porodica ukazuje da su nastali amplifikacijom od jedne (ili svega nekoliko) izvornih kopija, a koje su rezultat delecije u dužem elementu. U određenom broju slučajeva je teško, ako ne i nemoguće, direktno povezati neku MITE porodicu sa autonomnim transpozonom prisutnim unutar istog genoma (Feschotte i sur. 2003). Najčešće je nukleotidna sličnost između MITE-a i prvog najbližeg autonomnog elementa samo u TIR sljedovima. Za objašnjenje ovog paradoksa postoje dvije hipoteze. Prva predlaže da su neki MITE elementi nastali *de novo* nakon povoljnog pozicioniranja samih TIR sljedova ili sljedova koji nalikuju na TIR-ove autonomnih transpozona. Druga je mogućnost da su MITE elementi relikti svojedobne invazije transpozona čije su autonomne kopije izbrisane ili nisu doživjele fiksaciju unutar populacije (Feschotte i sur 2003).

1.4. Evolucijska važnost mobilnih elemenata

Eukariotski organizmi pokazuju veliku varijabilnost što se tiče zastupljenosti transponirajućih elemenata u svojim genomima. Oni zauzimaju npr. 22% genoma vrste *Drosophila melanogaster* (Kapitonov and Jurka 2003), 35% genoma riže (IRGSP 2005) i gotovo 50% genoma čovjeka (Lander i sur. 2001). Utvrđeno je da imaju i velik utjecaj na razlike u veličini

genoma, čak između blisko srodnih vrsta (Kidwell 2002). Čini se da su retrotranspozoni ti koji igraju ključnu ulogu u naglom povećanju i smanjenju veličine genoma. Ogromne razlike također postoje i u relativnoj zastupljenosti transpozona i retrotranspozona u pojedinim vrstama, neovisno o njihovim apsolutnim vrijednostima (Feschotte i Pritham 2007, Slika 1.5).



Slika 1.5: Relativna zastupljenost retrotranspozona i DNA transpozona u različitim eukariotskim genomima. Kratice označavaju sljedeće vrste - Sc: *Saccharomyces cerevisiae*; Sp: *Schizosaccharomyces pombe*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Os: *Oryza sativa*; Ce: *Caenorhabditis elegans*; Dm: *Drosophila melanogaster*; Ag: *Anopheles gambiae*; Aa: *Aedes aegypti*; Eh: *Entamoeba histolytica*; Ei: *Entamoeba invadens*; Tv: *Trichomonas vaginalis*. (Preuzeto iz Feschotte i Pritham 2007)

Zbog mnogih svojih karakteristika, transponirajući elementi ostavljaju dojam „parazitske“ DNA u genomu. Unatoč tome, brojni noviji radovi pokazuju kako transponirajući elementi imaju značajnu ulogu u evoluciji gena i strukture genoma eukariotskih organizama. Tako transponirajući elementi mogu imati aktivnost regulatornih faktora (Lister i sur. 1993, Yang i sur. 1998), utjecati na izrezivanje transkripcijskih produkata (Marillonnet i Wessler 1997), promovirati rearanžmane između eksona (Moran i sur. 1999), ili utjecati na cjelokupna kromosomska restrukturiranja (Végh i sur. 1990). Homologna rekombinacija među transpozonomima može biti uzrokom delecija, inverzija, translokacija i segmentalnih duplikacija, kao i mutacijskih promjena ukoliko transponiraju unutar gena (Belancio i sur. 2008). Zbog njihove mobilnosti, brojnosti i raznolikosti, ti elementi predstavljaju dragocjen doprinos evolucijskim procesima. Analiza baze humanih promotorskih sekvenci pokazuje da 25%

humanih promotora sadrži sljedove transpozonskog porijekla (van de Lagemaat i sur. 2003, Jordan i sur. 2003), dok 20% ljudskih i mišjih gena sadrže transpozone u svojim netranslatiranim regijama (van de Lagemaat i sur. 2003). Osim ranije spomenutog proteina CENP-B, postoji cijeli niz transkripcijskih faktora i DNA - vezujućih proteina koji vuku porijeklo od različitih nadporodica transpozaza (pregled u Feschotte 2008). Transponirajući elementi također igraju važnu ulogu u formiranju heterokromatina. Kod kvasca *Schizosaccharomyces pombe* RNAi-ovisni proces formiranja heterokromatina na području LTR sljedova može u vegetativnim stanicama uzrokovati represiju transkripcije gena specifičnih za mejozu (Schramke i Allshire 2003). Retrotranspozoni HeT-A i TART su zaslužni za održavanje heterokromatina na kromosomskim krajevima vrste *Drosophila* (Pardue i Debaryshe 2000). Što se tiče same centromere, već je ranije spomenuta prisutnost motiva CENP-B u centromernim ponavljanjima te potencijalno transpozonsko porijeklo proteina CENP-B, kao i prisutnost retrotranspozona L1 u sklopu humane centromere. Postojanje retrotranspozona specifičnih samo za centromerno područje također ide u prilog hipoteze o njihovoj mogućoj ulozi u izgradnji centromere. Primjer je retrotranspozona Ty3/gypsy koji se može naći kod različitih vrsta trava (Langdon i sur. 2000). Lokaliziran je isključivo na centromernom području, a u nukleotidnoj sekvenci je konzerviran oko 80% čak i među vrstama koje su divergirale prije više od 55 milijuna godina. Retrotranspozoni sličnih svojstava nađeni su i kod centromera žitarica, kukuruza, riže, ječma, pšenice i raži (Zhang i sur. 2004, Jiang i sur. 2003, Nagaki i sur. 2003). U prethodnim primjerima centromerna ponavljanja su potpuno zadržala strukturu transpozona, no postoje i one varijante koje zadržavaju sličnost s transpozonom samo u jednom dijelu sekvence uzastopnih ponavljanja. Primjer su 250 pb duga centromerna ponavljanja ječma *Aegilops speltoides* koja pokazuju visok stupanj homologije s dijelom retrotranspozona koji je nalik elementu Ty3/gypsy (Cheng i Murata 2003).

1.5. Povezanost satelitnih sekvenci i mobilnih elemenata

Sam nastanak satelitnih DNA, odnosno inicijalnih uzastopno ponovljenih monomernih nizova predstavlja područje koje još nije u potpunosti razjašnjeno. Primijećena je podudarnost nukleotidne sekvence nekih satelitnih DNA i mobilnih elemenata te se pretpostavlja da bi mobilni elementi mogli biti izvorište jedinica ponavljanja. Već je spomenuta sličnost između mobilnih elemenata i satelitnih DNA u centromernom području, no ta pojava nije ograničena

samo na centromere, već je prisutna u različitim dijelovima genoma i kod cijelog spektra različitih organizama. Tako satelitna DNA pvB370 nađena u skupini vrsta *Drosophila virilis* posjeduje značajnu sličnost u sekvenci s LTR sekvencama transpozona pDv (Heikkinen i sur. 1995), a sličan odnos primijećen je i kod satelitne DNA Sobo u krumpiru (Tek i sur. 2005). Satelitna DNA *HinfI* nađena u genomu pilića sadrži retrotranspozon CR1 u sklopu svojih ponavljanja (Li i Leung 2006). Kod pericentromerno locirane glavne satelitne DNA kitova reda *Cetacea* primijećeno je 63% sličnosti s 3' krajem elementa L1 (Kapitonov i sur. 1998). Slično tome, satelitna ponavljanja koja se nalaze u pericentromernom području vrste *Arabidopsis thaliana* pokazuju ~72 –79% sličnosti s 5' krajem DNA transpozona En/Spm (Kapitonov i Jurka 1999). Kod kornjaša *Tribolium castaneum* je uočeno da se sekvenca TCAST satelita pojavljuje u različitim organizacijskim oblicima: dugim nizovima uzastopno ponovljenih monomera, u obliku parcijalnih monomernih jedinica ili kratkih nizova (1 - 4 jedinice), strukturama nalik DNA transpozonu, te čak i u 5' netranslatiranoj regiji retrotranspozona CR1-3 TCa, što objašnjava mobilnost ove sekvence i činjenicu da se može pronaći i u eukromatinskom dijelu genoma, u neposrednoj blizini gena (Brajković i sur. 2012).

Zanimljivo je također izdvojiti i povezanost satelitnih DNA i MITE mobilnih elemenata. Potencijal MITE i njima srodnih elemenata da dovedu do stvaranja nizova satelitnih DNA primijećen je u vrstama roda *Drosophila* (Miller i sur. 2000), u rodu *Xenopus* (Hikosaka i Kawahara 2004) te u vrstama školjkaša (Gaffney i sur. 2003, Lopez-Flores i sur. 2004). U vrstama *Drosophila subobscura* i *D. madeirensis* tako su opisane SGM-IS raspršeno ponovljene sekvence koje su po strukturi srodne MITE elementima, dok je iz sekvenci SGM-IS proizašla i jedinica ponavljanja satelitne DNA u vrsti *D. guanche* (Miller i sur. 2000). Element CvA nađen je u vrsti *Crassostrea virginica*, a pripada porodici *pearl* elemenata koji sliče MITE elementima (Gaffney i sur. 2003). Podudarnost u nukleotidnoj sekvenci između uzastopnih ponavljanja koja tvore središnji dio mobilnog elementa i monomera satelitnih DNA prisutnih kod većeg broja školjkaša (Gaffney i sur. 2003, Lopez-Flores i sur. 2004, Plohl i sur. 2010) ukazuje na mogućnost da satelitne DNA različitih vrsta školjkaša imaju svoje porijeklo u ovom ili nekom sličnom mobilnom elementu.

Model nastanka satelitnih DNA iz MITE elemenata predložili su Hikosaka i Kawahara (2004) zahvaljujući istraživanju ponovljenih sekvenci u rodu *Xenopus*. U genomima tih vrsta nađene su dvije uzastopno ponovljene sekvence, Xstir (*X. leavis* i *X. borealis*) i tropXstir (*X. tropicalis*) pri čemu su obje slične različitim dijelovima MITE elementa Xmix (također prisutnom u tim vrstama). Zaključak je autora da su te dvije satelitne DNA nastale nezavisno iz MITE elementa Xmix i to nakon razdvajanja linija kojima pripadaju te vrste. Model koji predlaže kako dolazi do nastanka uzastopnih ponavljanja unutar samog MITE elementa uključuje stvaranje ukosnica između dva susjedna MITE elementa na jednolančanoj DNA (posljedica kašnjenja jednog lanca tijekom replikacije DNA), te izrezivanju dijela ukosnice djelovanjem nukleaza. Kao posljedica ovog procesa, novi MITE element će unutar sebe posjedovati dva puta uzastopno ponovljeni dio sekvence, a istim procesom može doći i do produljenja niza. Ti prvotni kratki nizovi još uvijek predstavljaju raspršeno ponovljene sekvence, no one bi se mogle, uz sudjelovanje već opisanih rekombinacijskih procesa, razviti u dugačke uzastopno ponovljene nizove, odnosno u nove satelitne DNA.

1.6. Osnovna obilježja školjkaša i njihovog genoma

Razred školjkaša (lat. bivalvia) se nalazi unutar carstva životinja, koljena mekušaca i obuhvaća oko 9200 vrsta koje se nalaze u morskim i slatkovodnim staništima cijeloga svijeta. Morski školjkaši obuhvaćaju 8000 vrsta smještenih u 99 porodica. Najveća je porodica Veneridae s više od 680 vrsta a slijede porodice Tellinidae i Lucinidae, svaka s 500 vrsta. Školjkaši imaju važnu ulogu u ekosustavima koje naseljavaju (Gosling 2003), a služe i kao potpora mnogim drugim zajednicama mekušaca, te sudjeluju u prijenosu minerala i organske tvari u bentoska staništa. Prilagođeni su mnogim staništima - od slatke vode, dubokomorskih područja slabe osvjetljenosti, do nemirnih i kisikom siromašnih staništa. Kao posljedica filtriranja tijekom hranjenja, mogu nakupiti različite ksenobiotske supstance pa su školjkaši organizmi od odabira u monitoringu okoliša. Nakupljanje ksenobionata istovremeno predstavlja i prijetnju ljudskom zdravlju prilikom korištenja u prehrambene svrhe. Uzevši u obzir značaj školjkaša u morskim ekosustavima i marikulturi, sukladno tome i interes za genetičkim istraživanjima na tim organizmima postaje sve veći. *NCBI GenBank* donosi informacije o genomskim projektima 3 vrste školjkaša *Pinctada margaritifera*, *Pinctada maxima* i *Crassostrea gigas*. Istraživanja na prve 2 vrste zasad pokrivaju područje transkriptoma i genske ekspresije školjkaša, a vrsti *Crassostrea gigas* je sekvenciran genom,

iako informacije o nukleotidnim sljedovima još nisu pridružene pojedinim kromosomima. U sklopu određivanja nukleotidnih sljedova genoma tog školjkaša napravljena su istraživanja o njegovoj prilagodbi na stresne uvjete okoliša i o kompleksnim procesima koji sudjeluju u formiranju ljuštore (Zhang i sur. 2012). Početni pokušaji sastavljanja nukleotidnih sljedova genoma te analize različitih tipova izoliranih sekvenci su provedene i za vrstu *Pinctada fucata*, školjku bisernicu (Takeuchi i sur. 2012).

Pregledni rad Saavedra i Bachère (2006) donosi informacije o tome koja su sve svojstva školjkaša do sad istraživana s genetičke perspektive. Prikupljene su neke informacije o genima: odgovornim za imunitet i otpornost na bolesti, prihvaćanju za podlogu i stvaranju ljuštore, otpornosti na onečišćenja i stres, za metaboličke procese, respiraciju, pohranu energije, metabolizam kalcija, probavu, razvoj, mišićnu kontrakciju, živčanu aktivnost i reprodukciju. Također, prikupljene su i informacije o ribosomalnim i mitohondrijalnim sekvencama DNA koje su potom korištene u filogenetskim i populacijsko-genetičkim studijama. Dosadašnja istraživanja na ovim organizmima uglavnom su rezultat zanimanja za točno određenim funkcionalnim sekvencama, dok novija istraživanja kreću u smjeru istraživanja cjelokupnog genoma. Genetičke analize svojstava koja su od važnosti za uzgoj školjkaša temelje se i na proučavanju njihovih transkripcijskih profila, pri čemu se biblioteke cDNA, baze EST i analize *microarray-a* koriste za otkrivanje gena uključenih u stanične i biokemijske procese zaslužne za ta svojstva. Rastući broj podataka za mekušce u GenBank EST bazama može poslužiti za komparaciju srodnih genoma i bolje razumijevanje njihove evolucije. Tanguy i sur. (2008) nam donose analizu 13 013 EST sekvenci školjkaša *Crassostrea gigas*, *Mytilus edulis*, *Ruditapes decussatus* i *Bathymodiolus azoricus*, dok u bazi podataka GenBank možemo pronaći 365 912 EST sekvenci porijeklom od 33 vrste školjkaša. Osim analize kodirajućih sekvenci, uzastopno i raspršeno ponovljene sekvence su također bile predmetom znanstvenih istraživanja kod tih organizama (detaljniji pregled kasnije). Na školjkašima je rađen i velik broj citogenetičkih studija pa tako Leitão i Chaves (2008) donose iscrpan pregled dosadašnjih citogenetičkih bojenja i fluorescencijskih hibridizacija *in situ* (eng. *Fluorescent in Situ Hybridization*, skraćeno FISH) provedenih na 51 vrsti školjkaša.

U sklopu ovog rada provedena su istraživanja na 3 vrste školjkaša: *Donax trunculus*, *Ruditapes decussatus* i *Ruditapes philippinarum*. *R. decussatus* i *R. philippinarum* pripadaju istom rodu porodice Veneridae, dok se *D. trunculus* nalazi unutar porodice Donacidae. Školjkaš *R. decussatus*, poznat pod nazivom: „kućica“, „vongola“, ili „lisanka“ (Slika 1.6a) autohtona je vrsta koja živi uzduž istočne obale Jadranskog mora te je vrlo cijenjena zbog izuzetno kvalitetnog i ukusnog mesa. U novije vrijeme ova vrsta postaje ozbiljno ugrožena od srodne vrste *Ruditapes philippinarum* (Slika 1.6b), koja je početkom 1980-tih godina unesena u Jadran iz jugoistočne Azije (Gosling 2003). Premda je vrsta *R. philippinarum* inicijalno nastanjena uz talijansku obalu u komercijalne svrhe, svojom invazivnošću i dobrom prilagodljivošću počela je potiskivati autohtonu vrstu *R. decussatus* (Jensen i sur. 2004), kolonizirajući njezina staništa i uz hrvatsku obalu Jadrana. Međutim, u usporedbi s autohtonom vrstom *R. decussatus*, unesena invazivna vrsta *R. philippinarum* slabije je kvalitete mesa te je time i manje cijenjena na tržištu. Morfološka sličnost školjkaša *R. decussatus* i *R. philippinarum*, zajedno sa stanovitom varijabilnosti u izgledu vanjske površine ljušture unutar svake od tih vrsta, te naseljavnje istih tipova staništa znatno otežavaju *in situ* i morfološko razlikovanje ovih školjaka. Povrh toga, nedavno su uz sjeverozapadnu obalu Španjolske pronađene jedinice za koje je molekularno-genetičkim metodama dokazano da su križanci između *R. decussatus* i *R. philippinarum* (Hurtado i sur. 2011). Vrsta *D. trunculus* (slika 1.6c) također je od komercijalnog interesa, osobito uz atlantsku obalu Pirinejskog poluotoka. Prisutna je u Jadranu, no kod nas nije od gospodarskog značaja.



Slika 1.6: Tri vrste školjkaša koji su predmet istraživanja ovog rada.

1.7. Obilježja i evolucija ponovljenih sekvenci DNA kod školjkaša

Nakon razgradnje genomske DNA restrikcijskom endonukleazom *ApaI*, otkrivene su tri različite satelitne DNA u vrsti *Mytilus edulis* (Mytilidae), a koje su kasnije utvrđene i u srodnim vrstama *Mytilus galloprovincialis*, *Mytilus trossulus* i *Mytilus californianus* (Ruiz-Lara i sur. 1992, Martinez-Lage i sur. 2002). Duljina monomera tih satelitnih sekvenci nazvanih 1, 2 i 3 iznosi 173 pb, 161 pb i 166 pb. Sva tri satelita su AT bogata (55-65%), a u vrsti *M. edulis* analiza sekvence pokazuje visok stupanj homologije među monomerima (91%). Utvrđivanje tih satelita i kod srodnih vrsta omogućilo je usporedbu vrsno specifičnih satelitnih profila kao i rekonstrukciju filogenetskih odnosa među njima. Filogenetska stabla konstruirana na osnovu tih satelitnih sekvenci pokazuju da je *M. californianus* najdivergentija vrsta što je u skladu s morfološkim opažanjima, kao i sa studijama zasnovanim na mitohondrijalnoj DNA (Martinez-Lage i sur. 2002). Količina pojedinih satelitnih sekvenci u genomu razlikuju se od vrste do vrste, no kod svih je satelit 2 najzastupljeniji i čini <1,66% genoma, odnosno prisutan je u 140 000 kopija, dok je satelit 3 najmanje zastupljen (< 0,01%, oko 1100 kopija). Lokalizacija ovih satelitnih DNA na kromosomima provedena je FISH analizom na vrstama *M. edulis*, *M. galloprovincialis* i *M. trossulus* (Martinez-Lage i sur. 2002). Dok je satelit 2 nasumično raspoređen u malim skupinama po svim kromosomima, sateliti 1 i 3 su smješteni u subteloemernim područjima pojedinih kromosoma.

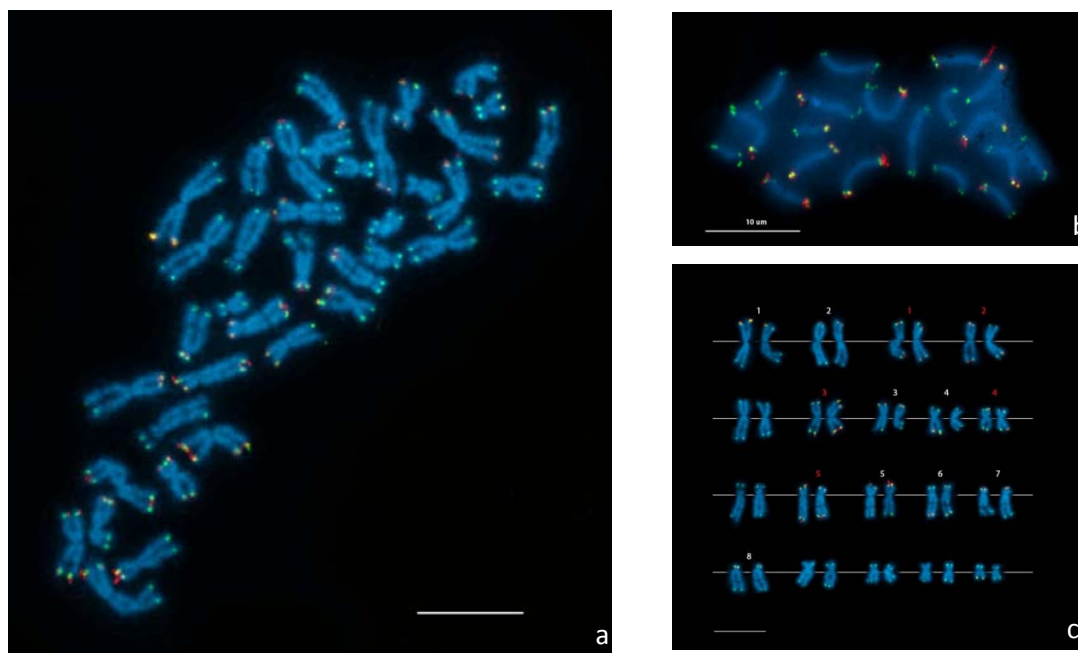
U vrsti *Adamussium colbecki* je nakon razgradnje genomske DNA endonukleazom *BglI* otkrivena satelitna DNA duljine monomera 170 pb (Canapa i sur. 2000). Analiza nukleotidnih sljedova pokazala je da se monomeri tog satelita mogu svrstati u nekoliko podporodica. AT sastav ove satelitne DNA je visok (59,2% - 73%), a unutar monomera postoje i brojni nizovi uzastopnih adenina ili timina. Osim toga, unutar monomera nađene su i sekvence slične CENP-B i CDEIII motivima, što znači da satelitna DNA posjeduje neke od značajki koje mogu utjecati na oblikovanje heterokromatina te organizaciju i aktivnost centromera. Kasnija istraživanja pokazala su da je ona uistinu smještena u centromernim područjima kod nekoliko kromosoma školjkaša *A. colbecki* (Odierna i sur. 2006).

U vrsti *Pecten maximus* (Pectinidae) utvrđeno ukupno šest različitih ponavljajućih sekvenci DNA (Biscotti i sur. 2007). Tri ponavljajuće DNA posjeduju duljinu monomera od oko 40 pb (porodice 2, 3 i 4), u sljedeće dvije monomeri su dugački 156 (porodica 5) odnosno 178 pb (porodica 1), a kod posljednje duljina monomera je 326 pb (porodica 6). Jedna od porodica s kratkom jedinicom ponavljanja vjerojatno se nalazi u sklopu jedinica ponavljanja višeg reda, dok druga porodica posjeduje unutarnje ponavljanje duljine 22 pb. Međusobna sličnost jedinica ponavljanja porodice 1 iznosi oko 90% a udio u genomu je malen, oko 0,13%. Monomeri porodice 5 izdvajaju se uzastopnim ponavljanjima jedne baze, i to naročito G, odnosno po ponavljanjima GG, AA, TT, CC i AGGG motiva. Najduža jedinica ponavljanja, ona porodice 6, sastoji se od dva uzastopna ponavljanja duljine 163 pb, čija međusobna sličnost iznosi 75%. Autori su utvrdili da ona posjeduje i određenu razinu sličnosti sa DTE satelitnom DNA (70% na 80 pb) iz vrste *Donax trunculus*.

U vrsti *D. trunculus* dosad je opisano sedam satelitnih DNA, što je čini najbolje istraženom vrstom školjkaša iz perspektive pregleda nad uzastopno ponovljenim sekvencama DNA u genomu. Osnovno obilježje satelitnog profila ove vrste je suživot brojnih divergentnih satelitnih podgrupa, prisutnih uglavnom u malom broju kopija. Četiri opisane satelitne DNA u vrsti *D. trunculus* pripadaju porodici DTHS otkrivenoj nakon cijepanja genomske DNA s endonukleazom *HindIII* (Plohl i Cornudella 1997). Duljine monomera pojedinih satelitnih DNA unutar porodice kreću se u rasponu od 136 pb do 167 pb. Sve četiri podporodice su niskozastupljene: DTHS1 čini 0,043% genoma i prisutna je u 3 600 kopija, DTHS2 0,081% (6 600 kopija), DTHS3 0,035% (3 300 kopija) dok DTHS4 čini 0,008% genoma (800 kopija). Usprkos razlici u sekvenci, članovi porodice DTHS imaju srodne oligonukleotidne motive ponovljene i do 15 puta unutar satelitnih monomera. Autori smatraju kako je upravo TTAGGG motiv osnovni građevni blok monomera nekih satelitnih DNA porodice *HindIII*. Ta heksamerna jedinica ponavljanja je karakteristična telomerna sekvenca prisutna kod kralješnjaka, a nađena je i na krajevima kromosoma kod vrste *D. trunculus* te nekih drugih morskih beskralješnjaka (Plohl i sur. 2002).

Nakon cijepanja genomske DNA vrste *D. trunculus* endonukleazom *HinfI* otkrivene su još dvije različite satelitne DNA, izdvojene iz iste elektroforetske pruge, kojima je jedina zajednička značajka duljina jedinice ponavljanja od 169 pb (Petrović i Plohl 2005). Jedan od ta dva satelita je DTF1 koji je AT bogat (62%) i izgrađuje 0,1 % genoma vrste *D. trunculus*, što znači da se u genomu nalazi oko 8200 kopija takvog monomera. Analizirani monomeri grupiraju se u dvije podskupine koje se međusobno razlikuju 17%, no unutar svake podskupine monomeri su izrazito konzervirani. Opisana je i jedna organizacijska jedinica višeg reda (DTRS), čiji su sastavni dijelovi jedan potpuni i jedan skraćeni DTF1 monomer te 14 bp dugačak nesrodni nukleotidni slijed. Dokazano je da je ta organizacijska jedinica također uzastopno ponovljena u genomu (Petrović i Plohl 2005). Analiza nukleotidne raznolikosti na razini jedinice ponavljanja pokazala je da postoji neravnomjerna raspodjela mutacija, odnosno da se u svim DTF1 monomerima izdvaja segment duljine 30 pb, kod kojeg je sačuvanost sekvence izrazito velika. Takva sačuvanost nukleotidnog slijeda može pokazivati evolucijski pritisak na dio monomera, što se kod satelitnih DNA dovodi u vezu s interakcijom s proteinima koji se vežu na DNA te ulogom u nastanku i održavanju heterokromatina (Petrović i Plohl 2005).

Druga spomenuta satelitna DNA je DTF2. Zanimljiva je zbog visokog udjela GC nukleotida u sekvenci monomera (62,5%), te udjela u genomu od 2% koji višestruko nadilazi uobičajene satelitne DNA školjkaša. Naime, kod većine dosad proučenih vrsta školjkaša udio pojedinih satelitnih DNA u genomu je u pravilu < 1 % (Martinez-Lage i sur. 2002, Muchmore i sur. 1998, Plohl i Cornudella 1996, 1997). Analiza nukleotidnog slijeda jedinica ponavljanja pokazala je da se radi o satelitnoj DNA izrazito sačuvane sekvence, i to jednoliko cijelom duljinom monomera. Uz 169 bp duge jedinice ponavljanja (DTF2L varijanta) pronađena je i kraća varijanta (DTF2S), duga 155 pb. DTF2 je ujedno i prva satelitna DNA vrste *D. trunculus* kojoj je utvrđen položaj na kromosomima. Ukupni broj kromosoma vrste *D. trunculus* iznosi $2n = 38$, a ova satelitna DNA je lokalizirana u subtelomernim područjima kromosoma. Pritom, kod sedam kromosomskih parova nalazi se na jednom kraju, a kod pet pari je smještena na oba kraja kromosoma, i to samo u nekim od područja za koja je bojanjem pokazano da su GC-bogata (Petrović i sur. 2009).



Slika 1.7: Kromosomi vrste *D. trunculus* na kojima je FISH analizom utvrđen položaj DTF2 satelitne DNA (crveni signali) i telomernih ponavljanja (zeleni signali). Žuti signali su rezultat blizine crvenih DTF2 signala i zelenih telomernih signala. a) metafazni kromosomi, b) kromosomi u mejozi, c) kariogram (preuzeto iz Petrović i sur. 2009).

DTE satelitna DNA, duljine monomera 155 pb, otkrivena je u vrsti *D. trunculus* nakon razgradnje s endonukleazom *EcoRV* (Plohl i Cornudella 1996). Ta satelitna DNA sadrži dvije skupine monomera, DTE1 i DTE2, čiji se nukleotidni sljedovi međusobno razlikuju ~11% i predstavljaju skupine koje se mogu jasno razlučiti preko dijagnostičkih mutacija. Udio DTE1 u genomu iznosi 0,23% (20 000 kopija), a DTE2 0,09%, odnosno 7600 kopija. Ova satelitna DNA pokazuje sličnost s dijelom satelitne DNA nađene u vrsti *Pecten maximus* (Biscotti i sur. 2007) te sa satelitnom DNA kamenica (Wang i sur. 2001, Lopez-Flores i sur. 2004) kao i s mobilnim elementom CvA (Gaffney i sur. 2003), koji će biti detaljnije opisan kasnije.

Clabby i sur. (1996) su otkrili satelitnu DNA duljine monomera 168 pb koja izgrađuje 1 - 4% genoma vrste *Crassostrea gigas* (Ostreidae). Nukleotidna sekvenca tog satelita je AT bogata (60%) a analiza pokazuje visok stupanj homologije među monomerima (79 - 94%). Kasnija

istraživanja su pokazala da su te satelitne sekvence smještene u centromernim područjima kromosoma. Isti satelit nije pronađen u srodnoj vrsti *Crassostrea virginica*, barem ne u količini koja bi se mogla utvrditi FISH analizom (Wang i sur 2001).

Ista satelitna DNA je pronađena i detaljnije opisana u sedam vrsta kamenica koje pripadaju rodovima *Ostrea* i *Crassostrea* nakon razgradnje restrikcijskom endonukleazom *HindIII* (Lopez-Flores i sur. 2004). Konsenzus duljina monomera u tim vrstama iznosi oko 166 pb, a potvrđen je visok AT sastav (63%), pri čemu su zapaženi dulji nizovi izgrađeni od A i T nukleotida. Osim s DTE satelitom otkrivenim u vrsti *D. trunculus* (Plohl i Cornudella 1996), primijećena je sličnost i s dijelom mobilnog elementa CvA nađenog u genomu vrste *Crassostrea virginica* (Gaffney i sur. 2003).

Posebno je zanimljiva satelitna DNA BIV160 koja je opisana kod 9 vrsta (*Glycymeris glycymeris*, *Mya arenaria*, *Dosinia exoleta*, *Venus verrucosa*, *Ruditapes pullastra*, *Ruditapes rhomboides*, *Ruditapes decussatus*, *Ruditapes philippinarum*, *Nucula* sp.) unutar 3 glavne skupine školjkaša (Protobranchia, Pteriomorphia i Heteroconchia). Procjenjuje se da je starosti oko 540 milijuna godina, što ju čini najstarijom satelitnom DNA opisanom do sada (Plohl i sur. 2010). Prosječni AT sastav te satelitne DNA je 61%, a duljine monomernih jedinica kreću se od 158 pb (kod *R. philippinarum*) do 166 pb (kod *V. verrucosa*). Monomeri izdvojeni iz 9 navedenih vrsta, nakon međusobne usporedbe, pokazuju 3 različita obrasca: oni koji se grupiraju kao vrsno specifični (*Mya arenaria*), oni koji se ne grupiraju kao vrsno specifični već se nalaze izmiješani sa monomerima svih ostalih vrsta (*R. decussatus*), te oni čiji monomeri se dijelom nalaze u prvoj, a dijelom u drugoj kategoriji (*R. philippinarum*, *R. rhomboides* i *R. pullastra*). Kod vrste *R. decussatus* duljina monomera satelitne DNA BIV160 je oko 166 pb, sličnost jedinica ponavljanja je u rasponu od 69 - 97% i izgrađuje 2% genoma. Kod vrste *R. philippinarum* izdvajaju se dvije skupine monomera – kraća varijanta s duljinom monomera 157 pb i dulja varijanta s duljinom monomera ~166 pb. Cjelokupni udio u genomu im je nizak, < 0,01%.

Dok je BIV160 jedina satelitna DNA zasad otkrivena u vrsti *Ruditapes decussatus*, u srodnoj vrsti *Ruditapes philippinarum* je od ranije poznata još jedna satelitna porodica. Unutar vrste *R. philippinarum* prethodno je opisana satelitna DNA BgIII400 (Passamonti i sur. 1998). Ta satelitna DNA je AT-bogata i ima 400 pb duge monomere unutar kojih se izdvajaju dvije podjedinice duljine 200 pb. Analiza sekvence pokazuje visok stupanj homologije među monomerima (oko 98%), a FISH analizom je pokazano da se ta satelitna DNA nalazi uglavnom u pericentromernim kromosomskim područjima. Njena prisutnost je utvrđena u vrstama *Venerupis aurea* i *Paphia undulata*, ali ne i u vrsti *Ruditapes decussatus*, što je, prema autorima, u skladu s rezultatima prethodnih analiza, prema kojima je vrsta *R. decussatus* srodnija upravo vrstama *V. aurea* i *P. undulata*, nego vrsti *R. philippinarum*.

Gaffney i sur. (2003) su proveli analizu genomske biblioteke vrsta *Crassostrea virginica* i *Anadara trapezia* te su otkrili više raspršeno ponovljenih elemenata: CvA, CvB, CvE, CvG i Ana-1. Na osnovu strukturnih značajki zaključili da se radi o neautonomnim mobilnim elementima, a novu porodicu elemenata nalik MITE elementima nazvali su *pearl*. Sličnost u nukleotidnoj sekvenci među elementima je vrlo mala, no pokazuju istu strukturnu organizaciju. Početni dio elemenata zauzima jedinstvena regija nakon koje slijedi niz (1 - 6) centralnih ponavljanja, pri čemu je prva jedinica ponavljanja u nizu gotovo uvijek krnja na 5' kraju. Centralna ponavljanja kod CvA, CvB, CvE i CvG su duljine od 150 – 180 pb, nakon čega slijede tetranukleotidna mikrosatelitna ponavljanja te ponovo jedinstvene regije (koje su unutar pojedine skupine elemenata konzervirane). Elementi na svojim krajevima posjeduju AA dinukleotide, kao posljedica duplikacija na mjestu insercije. CvA blizu krajeva sadrži subterminalna invertna ponavljanja od 11 pb, kao i dodatne invertno i direktno ponovljene motive, dok u sklopu svojih centralnih ponavljanja posjeduje box A, ali ne i box B promotorske sekvence RNA polimeraze III. CvA pokazuje sličnost u nukleotidnoj sekvenci s nekoliko prethodno opisanih porodica satelitnih DNA: satelitnom DNA HindIII prisutnom kod kamenica (Lopez-Flores i sur. 2004), satelitnom DNA DTE vrste *Donax trunculus* (Plohl i Cornudella 1996) te satelitnom DNA BIV160 nađenom kod različitih vrsta školjkaša (Plohl i sur. 2010).

Lopez-Flores i suradnici (2004) pretpostavljaju da su sekvence slične CvA elementu i satelitnoj DNA iz kamenica šire rasprostranjene među školjkašima. U vrsti *Mytilus galloprovincialis* otkrivena je Mg1 ponavljajuća sekvenca koja posjeduje značajnu sličnost s CvE elementom (Kourtidis i sur. 2007). Ta ponavljajuća sekvenca je jedna od tri otkrivene u intronima gena *hsc70* i *hsc71*. Kourtidis i suradnici su uočili kako pet puta ponovljeni monomer elementa Mg1 (duljine 169 pb), zajedno sa svojim okolnim područjima, pokazuju strukturnu sličnost s mobilnim elementom CvE. Time je bila dokazana prisutnost još jednog elementa iz porodice *pearl* (prvotno utvrđenog u kamenicama) i kod drugih vrsta školjkaša, što ostavlja otvorenu mogućnost za utvrđivanje široke rasprostranjenost ovih (i sličnih) mobilnih elementa, kao i s njima povezanih sekvenci satelitnih DNA.

1.8. Cilj

Cilj ovog istraživanja je stvoriti sliku o genomu odabranih vrsta školjkaša s obzirom na zastupljenost i strukturu nekodirajućih ponovljenih sekvenci DNA. Također, cilj je istražiti različite oblike ponovljenih sekvenci i njihove međusobne odnose te evolucijske puteve i mehanizme koji dovode do širenja ovih sekvenci i uzrokuju prevladavanje njihovih pojedinih oblika. Karakterizacija nekodirajućeg dijela genoma je provedena kod 3 vrste školjkaša od ekološke i komercijalne važnosti: *Donax trunculus*, *Ruditapes decussatus* i *Ruditapes philippinarum*.

U tu svrhu korištene su molekularno-genetičke metode, počevši od konstrukcije genomskih biblioteka ovih vrsta koje su u potrazi za različitim tipovima ponovljenih sekvenci pretražene hibridizacijskim sondama, određivanja nukleotidnih sljedova odabranih kloniranih fragmenata genomske DNA, utvrđivanja prisutnosti ciljnih sekvenci u genomu pomoću lančane reakcije polimerazom, određivanje organizacije i udjela sljedova od interesa putem Southern hibridizacije, do bioinformatičkih računalnih analiza na osnovu kojih su utvrđeni odnosi među dobivenim nukleotidnim sljedovima.

Analiza zastupljenosti i sastava ovih sljedova važna je za sveukupno razumijevanje strukture i funkcije genoma školjkaša, a ujedno omogućuje i bolji uvid u zastupljenost, dinamiku nastanka i širenja kako novih tako i već ranije poznatih oblika uzastopno i raspršeno ponovljenih sekvenci te bolje razumijevanje molekularnih mehanizama odgovornih za njihovu evoluciju.

2. Materijali i metode

2.1. Izolacija genomske DNA školjkaša

Genomska DNA školjkaša *Donax trunculus* izolirana je u laboratoriju prof. dr.sc. Juana Pasantes (University of Vigo, Španjolska). Genomske DNA ostala dva školjkaša *Ruditapes decussatus* i *Ruditapes philippinarum* izolirane su prema sljedećem protokolu. Od tkiva školjkaša pohranjenog u 100%-tnom etanolu uzeti su plašt, mišići i dio sifona, žiletom izrezani na manje komade te usitnjeni u tarioniku uz dodatak 8 ml homegenizacijskog pufera (160 mM saharoza, 80 mM EDTA, 100 mM Tris-HCl pH 8) u koji je dodano 400 µl 0.5%-tnog SDS-a i 60 µl proteinaze K (150 µg/ml), nakon čega je provedena razgradnja proteina preko noći na 65°C. Uslijedila je ekstrakcija fenol:kloroformom (1:1) te ekstrakcija kloroformom, ponovljena 2 puta. Odvajanje faza provedeno je centrifugiranjem u hladenoj (4°C) centrifugi (12000 rpm/12 min). DNA je zatim pretaložena dodavanjem 2.4 ml 1.2 M K-acetata centrifugiranjem 12000 rpm/20 min, pa dodavanjem apsolutnog etanola volumena jednakog uzorku. Nakon laganog miješanja uzorak je ostavljen 15 min na ledu i ponovljeno je centrifugiranje 12000 rpm/20 min. Talog je osušen i otopljen u 500 µl TE pufera uz dodatak RNaze A. Razgradnja RNazom provedena je inkubacijom na sobnoj temperaturi kroz 10 min. DNA je ponovo pretaložena dodatkom 34 µl 0.2 M Na-acetata i jednakog volumena apsolutnog etanola, talog dobiven centrifugiranjem (12000 rpm/12 min) je ostavljen da se osuši na sobnoj temperaturi i ponovo ispran dodatkom 70%-tnog etanola nakon čega je uslijedilo finalno centrifugiranje 12000 rpm/5 min. Osušeni talog je otopljen u odgovarajućem volumenu TE pufera a koncentracija izolirane genomske DNA procijenjena je elektroforezom na 1%-tnom agaroznom gelu i *Nanodrop* uređajem.

2.2. Razgradnja genomske DNA restrikcijskim endonukleazama

Reakcije razgradnje genomske i plazmidne DNA restrikcijskim enzimima izvedene su pod uvjetima specifičnima za pojedini enzim, prema uputama proizvođača. Količine enzima korištenih za razgradnju i vrijeme inkubacije varirali su ovisno o količini DNA i željenom stupnju razgradnje.

2.3. Umnažanje ciljnih nukleotidnih sljedova uz korištenje lančane reakcije polimerazom

Lančana reakcija polimerazom (eng. *Polymerase chain reaction*, skraćeno PCR) korištena je za utvrđivanje prisutnosti nukleotidnih sekvenci od interesa u genomima školjkaša.

2.3.1. Umnažanje Dt-BIV160 skupine sekvenci

U svrhu izolacije dodatnih jedinica ponavljanja satelitne DNA BIV160 u školjkašu *Donax trunculus* konstruirane su specifične početnice:

Naziv početnice	Nukleotidni slijed (5'-3')
bivF	TACATAGACTTATATAGGGAAAATC
bivR	TTTGACCCCAGGGGAATAATT

Reakcijska smjesa je sadržavala:

Reagens	Volumen (µl)
5 × Go Taq Flexi pufer	20
MgCl ₂ (25 mM)	16
dNTP mix (10 mM svaki)	8
bivF početnice (10 pmol/µl)	4
bivR početnice (10 pmol/µl)	4
Taq DNA polimeraza (5 u/µL)	1,25
DNA kalup (10 ng/µl)	1
H ₂ O	45,75
Ukupni volumen	100

Uvjeti umnažanja:

Program reakcije sastojao se od početne 5-minutne denaturacije na 94°C, te 30 ciklusa umnažanja koja se sastoje od tri koraka. Na kraju slijedi produžena reakcija sinteze na 72°C kroz 7 minuta da bi se osigurao završetak sinteze svih započetih lanaca.

Stadij ciklusa	Temperatura i trajanje
Denaturacija	95°C (30")
Sparivanje klica i kalupa	55°C (30")
Sinteza DNA	72°C (30")

2.3.2. Dobivanje cjelovite sekvence dugih ukloniranih genomskih fragmenata

Zbog veličine pojedinih genomskih fragmenata i nemogućnosti dobivanja njihovog kompletnog nukleotidnog slijeda sekvenciranjem uz korištenje samo plazmidnih početnica, za pojedne klonove od interesa kod sva 3 školjkaša konstruirane su i dodatne početnice:

Naziv početnice	Nukleotidni slijed (5'-3')
DTC34AluF	AACCCAAGGATTAGGATAGAT
DTC34AluR	ACGCGTTTGTGGTCAG
P18F	CAGCATTGGTTTGTTCATCTTAT
P18R	AATAATTTGTTGAGTAGGTGGTCT
P18Fprolong	TGCTACGTTCCCAAGTTTCAC
P18Rprolong	CAA AATTGGCCCCGTCTCG
PD53F	GACCTCGTGACTTACTTTCTTGT
PD53R	CTCCTGTGACCTAGTTTTTGACCA
DTC9AluF	CAAACGGCTGATAGAACATA
DTC9AluR	TCACAAAGGACAATCATAACACC
DTC9AluFprolong	ATTCTAAACAACATAAACGAGTAA
DTC9AluRprolong	TTCTGAAAGGCGTCTGTAGTTA
DTC33AluF	AGGGGTGCACTACTTATG
DTC33AluR	ATATACATCTTTACCTCACGA
DTC84AluF	TTTGTTATGGAGTTATTCACCTAT
DTC84AluR	GTGTATAAACGTTGTAAGTGCTAA
D14F	ATAACAGTACGATTTCTTTCAA
D14R	AAGAGATAATGACATACAGAATAG
P46F	ATAAAAATAAAGAAAGAAAACC
P46R	ATGAACAGCGGCGTATG

2.3.3. Umnažanje centralnih ponavljanja DTC84

U svrhu daljnje analize DTC84 elementa u genomu školjkaša *D. trunculus* konstruiran je set početnica specifičan za umnažanje centralnih ponavljanja.

Početnice:

Naziv početnice	Nukleotidni slijed (5'-3')
DTC84AluSatF	TTGCCTGTGACGTCTACTTGTGC
DTC84AluSatR	AGAGGTCACAGGCAACCATCCA

Reakcijska smjesa je sadržavala:

Reagens	Volumen (μl)
GoTaq Green Master Mix*	25
DTC84AluSatF početnice (10 pmol/μl)	2
DTC84AluSatR početnice (10 pmol/μl)	2
DNA kalup (50 ng/μl)	1
H ₂ O	20
Ukupni volumen	50

* GoTaq DNA polimeraza u 2 × Green GoTaq puferu (pH 8.5), 400 μl dATP, 400 μl dGTP, 400 μl dCTP, 400 μl dTTP i 3 mM MgCl₂

Uvjeti umnažanja:

Program reakcije sastojao se od početne 5-minutne denaturacije na 94°C, te 35 ciklusa umnažanja koja se sastoje od tri koraka. Na kraju slijedi produžena reakcija sinteze na 72°C kroz 7 minuta da bi se osigurao završetak sinteze svih započetih lanaca.

Stadij ciklusa	Temperatura i trajanje
Denaturacija	95°C (30")
Sparivanje klica i kalupa	57°C (30")
Sinteza DNA	72°C (30")

2.3.4. Detekcija i karakterizacija dodatnih elemenata DTC84

Početicama specifičnim za konzervirane blokove koje okružuju centralna ponavljanja DTC84 elementa izvojene su dodatni primjerci, s različitim brojem centralnih ponavljanja.

Početnice:

Naziv početnice	Nukleotidni slijed (5'-3')
DTC84mobF	AACAAGAGCACCGCTGGGCG
DTC84mobR	CGCACGTTTGAAAAACGGGACGTA

Reakcijska smjesa je sadržavala:

Reagens	Volumen (μl)
5 × Go Taq Flexi pufer	4
MgCl ₂ (25 mM)	2
dNTP mix (10 mM svaki)	0,4
DTC84mobF početnice (10 pmol/μl)	0,8
DTC84mobR početnice (10 pmol/μl)	0,8
Taq DNA polimeraza (5 u/μL)	0,2
DNA kalup (10 ng/μl)	2
H ₂ O	9,8
Ukupni volumen	20

Uvjeti umnažanja:

Program reakcije sastojao se od početne 5-minutne denaturacije na 94°C, te 35 ciklusa umnažanja koja se sastoje od tri koraka. Na kraju slijedi produžena reakcija sinteze na 72°C kroz 7 minuta da bi se osigurao završetak sinteze svih započetih lanaca.

Stadij ciklusa	Temperatura i trajanje
Denaturacija	95°C (30")
Sparivanje klica i kalupa	55°C (30")
Sinteza DNA	72°C (1')

2.4. Pročišćavanje PCR produkata

Produkti PCR reakcija su pročišćeni na kolonama uputama iz kompleta "QIAquick PCR Purification Kit" (Qiagen). Fragmenti DNA su isprani s kolone Tris puferom (pH 8.5), u kojem su i pohranjeni.

2.5. Elektroforetsko odjeljivanje fragmenata DNA

Ovisno o tipu eksperimenta, DNA fragmenti su razdvojeni elektroforezom na agaroznom gelu različitih koncentracija (0.7-1.5%) u TAE puferu (40 mM Tris, 20 mM octena kiselina, 1 mM EDTA) uz dodatak etidij bromida (0.5 µg/ml). Trajanje elektroforeze ovisno o tipu uređaja, koncentraciji agaroze i veličini fragmenata DNA variralo je između 30 do 240 min, a fragmenti su vizualizirani pod UV-svjetlošću valne duljine 320 nm. Veličina fragmenata je utvrđivana u usporedbi s markerom „O'GeneRuler DNA Ladder Mix, ready-to-use“ (Fermentas), raspona veličina 100 - 10 000 pb.

2.6. Izolacija fragmenata DNA iz agaroznog gela

Komadići gela koji sadrže željene DNA fragmente izrezani su iz gela skalpelom te je provedena izolacija na kolonama prema uputama iz kompleta "QIAquick Gel Extraction Kit" (Qiagen). Fragmenti DNA su sa kolone isprani u Tris puferu (pH 8.5), u kojem su i pohranjeni. Nakon izolacije fragmenti su najčešće korišteni u svrhu kloniranja u plazmidni vektor ili za obilježavanje hibridizacijskih DNA sonda.

2.7. Ligacija fragmenata DNA u plazmidne vektore

Dobiveni DNA fragmenti od interesa, ovisno o tipu eksperimenta, ugrađeni su u različite vektorske sustave. PCR produkti su ligirani u pGEM-T Easy vektor pomoću enzima T4 DNA ligaze, sastavni dio kompleta „pGEM-T Easy Vector System” (Promega). Ligacijska smjesa inkubirana je 18 sati na 4°C.

Ostali DNA fragmenti ugrađeni su u dvije varijante pUC19 plazmidnog vektora koji je unaprijed bio pocijepan sa *Sma*I i *Bam*HI restrikcijskim enzimima i defosforiliran, te u takvom obliku komercijalno dostupan. Ligacija je provedena pomoću „Quick Ligation Kit“ (New England BioLabs) sustava na sobnoj temperaturi u trajanju 5 minuta.

2.8. Transformacija bakterije *Escherichia coli*

Tokom istraživanja su korišteni komercijalni sojevi kemijski kompetentnih "Subcloning Efficiency DH5 α Chemically Competent Cells" (Invitrogen) i elektrokompetentnih „ElectroMAX Stbl4 Competent Cells“ (Invitrogen) bakterijskih stanica. Transformacija je provedena prema uputama proizvođača, ovisno o tipu stanica.

2.9. Uzgoj bakterijskih stanica

Bakterijske stanice su uzgajane preko noći u tekućoj ili na krutoj hranjivoj podlozi. Tekuća hranjiva podloga je sadržavala (po litri): 10 g triptonu, 5 g kvašćevog ekstrakta, 5 g NaCl, uz dodatak penicilina (finalne koncentracije 100 μ g/ml). Kruta hranjiva podloga bila je istog sastava uz dodatak 15 g agara/l. Za plavo-bijelu selekciju klonova korišteni su IPTG (isopropyl β -D-1-thiogalactopyranoside) kao induktor i X-Gal (5-bromo-4-chloro-3-indolyl- β -

d-galactopyranoside) kao supstrat, nanešeno je po 50 μ l 100 mM IPTG i 50 μ l X-Gal (20 mg/ml) po ploči. Uzgoj je trajao preko noći na 37 ili 30°C, ovisno o tipu bakterijskih stanica i uputama proizvođača.

2.10. Izolacija plazmidne DNA

Plazmidna DNA izolirana je preko kolona za izolaciju i pročišćavanje plazmidne DNA iz nekoliko različitih komercijalno dostupnih kompleta: "QIAprep Spin Miniprep Kit" (Qiagen), „High Pure Plasmid Isolation Kit“ (Roche), „DNA-spin Plasmid DNA Purification Kit“ (Intron), prema uputama proizvođača. Koncentracija izolirane plazmidne DNA procijenjena je elektroforezom na 1%-tnom agaroznom gelu i/ili *Nanodrop* uređajem.

2.11. Određivanje primarne strukture DNA

Slijed nukleotida kloniranih DNA fragmenata određen je u servisima ustanove *Macrogen* (u Korei i Nizozemskoj) na ABI3730XL DNA Analyzer uređaju. Prilikom određivanja nukleotidnog slijeda korištene su komercijalno dostupne plazmidne početnice ili one specifično konstruirane za pojedine uzorke.

2.12. Obilježavanje hibridizacijskih sonda

Prilikom obilježavanja DNA probi digoksinom korištena je metoda nasumičnih početnica i komplet "DIG-DNA Labeling and Detection Kit" (Roche).

Uzorak DNA (~250 ng željenog fragmenta) denaturiran je 10 min pri 100°C i naglo ohlađen u ledu. Uzorku DNA (15 μ l) dodano je 2 μ l smjese nasumičnih heksanukleotida (1 μ g/ μ l), 1 μ l (2 U) Klenowljevog fragmenta, te 2 μ l dNTP smjese (uključujući DIG-11-dUTP). Inkubacija je trajala 18 sati pri 37°C a obilježena DNA pročišćena je taloženjem te otopljena u TE pufetu. Količina obilježene DNA određena je metodom kapljičnog testa, odnosno međusobnom usporedbom serije razrjeđenja novoobilježene sonde i kontrolne DNA poznatog stupnja obilježenosti.

2.13. Hibridizacijska analiza prema Southernu

Nakon elektroforeze u agaroznom gelu molekule DNA su prenesene na pozitivno nabijenu najlonsku membranu lužnatim kapilarnim prijenosom (u 0.4 M NaOH), čime je omogućena istovremena denaturacija DNA. Nakon prekonocnog prijenosa membrana je isprana u $2 \times$ SSC puferu i osušena, a DNA je zatim vezana na membranu inkubacijom 30 min na 120°C .

Usljedila je inkubacija u predhibridizacijskoj otopini (0.25 M Na_2HPO_4 pH 7.2, 1 mM EDTA, 20% SDS, 0.5% blokirajući reagens) 45 min – 2 h, na istoj temperaturi na kojoj će se kasnije nastaviti hibridizacija. Predhibridizacijska otopina je zatim zamijenjena svježom, u koju je dodana obilježena sonda DNA prethodno denaturirana kuhanjem 10 min na 100°C . Nakon denaturacije sonda je naglo ohlađena u ledu i dodana do konačne koncentracije od 10 ng/ml. Hibridizacija se nastavila preko noći na temperaturi koja ovisi o stupnju sličnosti traženih sekvenci i obilježene sonde kao i o željenoj selektivnosti hibridizacije (najčešće je bila u rasponu između 58 i 68°C). Nakon hibridizacije membrane su ispirane 3×20 min u puferu za ispiranje (20 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 1% SDS) na temperaturi 3°C nižoj od temperature hibridizacije.

Detekcija hibridizacijskih signala metodom kemiluminiscencije provedena je na sljedeći način. Sve reakcije su se odvijale na sobnoj temperaturi uz laganu trešnju. Membrane su inkubirane 5 min u puferu I (0.1 M maleinska kiselina, 3 M NaCl, 0.3% Tween 20, pH 8.0). Slijede inkubacije od 30 min u puferu II (1% blokirajući reagens u puferu I), te 30 min u puferu II uz dodatak digoksigeninskih antitijela (1:20 000) na koja je vezana alkalna fosfataza. Nevezano antitijelo se ispiru 5×10 min u puferu I, zatim je membrana inkubirana 2×5 min u puferu III (0.1 M Tris-HCl pH 9.5, 0.1 M NaCl). Na membranu je nakapan CDP-Star, kemilumiscentni supstrat za alkalnu fosfatazu razrijeđen 1:100 u puferu III nakon čega je membrana zataljena u plastičnu vrećicu. Nakon par minuta inkubacije, kad se reakcija počela razvijati, na membranu u vrećici se položio autoradiografski film. Duljina izlaganja filma je ovisila o jakosti hibridizacijskog signala a može trajati od par sekundi do nekoliko sati, po potrebi i preko noći.

2.14. Kapljična analiza DNA (eng. *dot blot*)

Određivanje udjela određenog nukleotidnog slijeda DNA u ukupnoj genomskoj DNA školjkaša provedeno je kapljičnom analizom. Metoda se temelji na hibridizaciji serije razrjeđenja genomske DNA, u prisutnosti serije razrjeđenja pozitivne i negativne kontrole, sa obilježenim nukleotidnim fragmentima od interesa. Serije razrjeđenja su pripremljene u TE puferu. Sva razrjeđenja denaturirana su zagrijavanjem 5 min na 100°C, i naglo ohlađena u ledu te zatim pomiješana s 200 µl 20 × SSC pufera. Ukupni finalni volumen svih razrjeđenja je nanešen na najlonsku membranu pomoću vakuumske uređaja. Membrana je zatim osušena a DNA vezana na membranu inkubacijom 30 min na 120°C. DNA na membrani je denaturirana u 0,4M NaOH 10 min, zatim neutralizirana ispiranjem u 1.5 M NaCl, 0.5 M Tris-HCl (pH 7.2) i osušena na zraku.

Hibridizacija membrane i detekcija signala provedeni su kao što je opisano u prethodnom poglavlju. Autoradigrafski filmovi prebačeni su u digitalni oblik pomoću skenera, a određivanje količine DNA od interesa provedeno je denzitometrijskom usporedbom jačine signala genomskih i kontrolnih razrjeđenja.

2.15. Analiza otisaka kolonija (eng. *colony lift*)

Bakterijske kolonije su uzgojene preko noći na krutoj hranjivoj podlozi uz prisutstvo antibiotika. Kružne membrane (promjera 83 mm) su na nekoliko mjesta asimetrično obilježene kako bi se u kasnijim postupcima mogla ustanoviti njihova izvorna orijentacija na ploči te pažljivo prislonjene na površinu bakterijskih ploča. U nastavku, membrane su položene na filter papir natopljen denaturacijskim puferom (0.5 M NaOH, 1.5 M NaCl) i inkubirane 5 min. Na isti način membrane su lagano kratko prislonjene na suhi filter papir pa pažljivo prebačene na novi, ovoga puta natopljen neutralizacijskim puferom (0.5 M Tris-HCl pH 7.4, 1.5 M NaCl) i inkubirane 3 min. Nakon toga su prebačene u Petrijeve zdjelice s 2 × SSC puferom pa su kroz nekoliko izmjena pufera, uz trešnju, uklonjeni s membrana svi tragovi bakterijskog materijala. Membrane su osušene na zraku a fiksiranje DNA na membrane je provedeno na 80°C tijekom 2 sata.

Bakterijske ploče s kojih je prebačen materijal na membrane su dodatno inkubirane 3 - 5 sati na 37°C kako bi kolonije ponovo narasle na ploči.

Hibridizacija membrane i otkrivanje signala provedeni su kao što je opisano u poglavlju *Hibridizacijska analiza prema Southernu*.

2.16. Računalne analize nukleotidnih sekvenci

Početna obrada sekvenci, sravnjivanje, obilježavanje pojedinih segmenata, konstrukcija početnica, analiza restrikcijskih mjesta, grafičko *dot-plot* uspoređivanje dvaju sekvenci te određivanje konsenzus sekvence za pojedine skupine sekvenci je provedeno programom *Geneious 5.4.3 (Biomatters Ltd.)*.

Filogenetske analize su provedene programom *Network* (Fluxus Technology Ltd 1999-2012).

Pretraživanja sekvenci u odnosu na substrukture, ponavljanja i različite motive su provedene programima *einverted*, *palindrome*, *equicktandem*, koji su sastavni dio EMBOSS kompleta dostupnog na mreži.

Za otkrivanje nukleotidnih slijedova koji su porijeklom od tRNA je korišten *tRNAscan-SE 1.21* servis.

Sve nove sekvence su uspoređene s onima u *NCBI GenBank* banci podataka preko BLAST servisa (Altschul i sur. 1990) i s bazom podataka *Repbase* (Jurka 2005), u čijem sklopu se nalazi CENSOR pretraživač za uspoređivanje dobivenih sekvenci sa kolekcijom različitih tipova poznatih repetitivnih sekvenci i mobilnih elemenata.

3. Rezultati

3.1. Pregled repetitivnih sekvenci u genomu školjkaša *Donax trunculus*

U svrhu detekcije i karakterizacije repetitivnih sekvenci genoma školjkaša *Donax trunculus*, napravljena je analiza otisaka kolonija (eng. *colony lift*) uz parcijalno pocijepanu, kompletnu genomsku DNA korištenu kao hibridizacijsku sondu. Taj pristup temelji se na činjenici da će repetitivne sekvence koje su najčešće prisutne u većem broju kopija u genomu dati intenzivnije hibridizacijske signale od nukleotidnih sekvenci prisutnih u samo jednoj kopiji. Hibridizacijom uzorka od oko 1000 kolonija koje sadržavaju fragmente genomske DNA školjkaša *D. trunculus* dobivene cijepanjem endonukleazom *AluI*, inicijalno je ovom metodom izdvojeno 67 kolonija. Nakon izolacije i utvrđivanja nukleotidnog slijeda, dobiveni genetički fragmenti od interesa uspoređeni su s bazom podataka *NCBI GenBank* te putem *CENSOR* pretraživača i s bazom *Repbase* koja sadrži zbirku različitih tipova poznatih repetitivnih sekvenci i mobilnih elemenata. Dobiveni rezultati prikazani su u Tablici 3.1.

Veličina haploidnog genoma ovog školjkaša iznosi $3,12 \times 10^9$ pb prema *Animal Genome Size Database*, pri čemu sekvencirani klonovi od interesa obuhvaćaju oko 160 000 pb. Duljina kloniranih fragmenata je u rasponu od 41 pb – 9200 pb, a njihov prosječan GC sastav iznosi 35.8%. Dobiveni sljedovi se mogu kategorizirati u 7 skupina s obzirom na različite tipove repetitivne DNA koje sadrže.

Prva skupina obuhvaća genomske fragmente koji sadrže isključivo uzastopno ponovljene sekvence (DTC4Alu, DTC50Alu, DTC52Alu). Druga skupina sadrži uzastopno ponovljene sekvence vezane uz nerepetitivne i zasad neokarakterizirane sljedove genoma školjkaša (DTC51Alu). Za nekoliko klonova iz te skupine je utvrđeno da sadrže dijelove nekih *D. trunculus* satelitnih DNA poznatih od ranije, prisutnih u genomu u malom broju kopija: DTC34Alu i DTC36Alu (DTHS1, Plohl i Cornudella 1997), DTC32Alu (DTHS3, Plohl i Cornudella 1997) te DTC38Alu (DTE, Plohl i Cornudella 1996). Sljedeće 3 grupe fragmenata su pokazale nukleotidnu sličnost s različitim tipovima mobilnih elemenata. S različitim porodicama DNA transpozona sličnost pokazuju sekvence DNA u genomskim fragmentima DTC7Alu, DTC8AluF, DTC12Alu, DTC14Alu, DTC17AluR, DTC20Alu, DTC21Alu, DTC30AluR, DTC37AluF, DTC42AluR, DTC45Alu, DTC47AluF, DTC63Dra, DTC70AluF, DTC73AluF, DTC75AluF, DTC81AluR, DTC85Alu, DTC3BamF, DTC3BamR. Sličnost s različitim porodicama LTR retrotranspozona imaju klonovi

DTC39Alu i DTC7BamF, dok sličnost s različitim porodicama non-LTR retrotranspozona postoji u DTC15AluF, DTC17AluF, DTC27AluR, DTC42AluF, DTC82AluF i DTC1BamR.

klon	Satelitna DNA		Transpozon			cDNA	ostalo	Duljina klona
	Samo satelitna DNA	Satelitna DNA prisutna uz neku drugu sekvencu	DNA transpozon	Retrotranspozon				
				LTR	Non-LTR			
DTC1Alu							407-588 <i>Macoma balthica</i> microsat.	635
DTC2Alu						664-778, 1433-1601	1446-1566 5S_CPo (raspršeno ponavljanje)	1741
DTC3Alu						664-778, 1433-1601	1446-1566 5S_CPo (raspršeno ponavljanje)	1741
DTC4Alu	BIV160 sat					1-375		397
DTC5Alu								653
DTC6Alu						194-315, 491-769		1155
DTC7Alu			1228-1279 Hovi2			552-629, 1477-1853	547-627, 1527-1598, 1801-1865 <i>Corbicula fluminea</i> Msx gen za Msx protein	2307
DTC8AluF			646-729 RAS2_MT					920
DTC8AluR						288-392, 461-571, 765-895	48-112 TguERV4N1_I Endogeni retrovirus	905
DTC9Alu			2577-2623 CR1-75_HM 2863-3019 Helitron-N3_SP 3467-3531 HELITRON7_OS	564-761 Gypsy123-I_DR		554-752, 2861-3040, 2938-3196	2871-3198 <i>Mercenaria mercenaria</i> microsat. 570-722 transponirajući genetički element 297 iz <i>Drosophila</i> (reverzna transkriptaza)	3609
DTC10Alu								215
DTC11AluF								921
DTC11AluR							618-662 ERV3-1_MEu-I	906

							Endogeni retrovirus	
DTC12Alu			7-186 CVE			2-92, 99-186	99-173 <i>Aulacomya atra</i> microsat.	190
DTC13AluF								921
DTC13AluR						151-359, 417-634, 701-749	133-191, 408-465, 699-743 <i>Corbicula fluminea</i> Msx gen za Msx protein	906
DTC14Alu			268-343 Helitron-1B_SBi					1680
DTC15AluF					171-237 CR1-41_HM			821
DTC15AluR								861
DTC16Alu								62
DTC17AluF		6-55 BIV160 sat			450-504 CR1-5_NV			870
DTC17AluR			123-234 HELITRON7_OS					862
DTC18Alu								1074
DTC19Alu								1121
DTC20Alu			842-890 hAT-45_HM					1000
DTC21Alu			97-144 TURMOIL1					1239
DTC22Alu						297-496		502
DTC23AluF								871
DTC23AluR								859
DTC24Alu								534
DTC25Alu							243-307 TguERV4N1_I Endogeni retrovirus	350
DTC26Alu						320-384, 443-690	344-423, 618-655 <i>Corbicula fluminea</i> Msx gen za Msx protein	812
DTC27AluF								871
DTC27AluR					445-491 L1-1B_Cpo	15-90, 158-256, 472-580, 527-624		859

DTC28Alu					47-144 CR1-32_HM 171-467 CR1-16_DR		47-104 Salmo salar zonadhesinu-sličan gen 52-486 reverzna transkriptaza mobilnog elementa jockey (Drosophila)	1311
DTC29Alu								65
DTC30AluF			163-267 hAT-40_HM 549-660 EnSpm1_TM					671
DTC30AluR								909
DTC32Alu		874-1453 DTHS3			370-487 CR1-21_HM	874-1456	878-1456 Spisula solidissima gen za zinc finger protein 230-550 reverzna transkriptaza mobilnog elementa jockey (Drosophila)	1456
DTC33Alu				813-863 ROVER_DM 2383-2432 BEL11-LTR_DR	322-395 CR1-3_TCa	1314-1411, 1629-1750	1935-1999 TguERV4N1_I Endogeni retrovirus	2711
DTC34Alu		1-57 DTHS1	783-828 MIRAGE1		19-56 Coprina_Cc1			1406
DTC35AluF								822
DTC35AluR						566-652	566-649 Spisula solidissima nHb	810
DTC36Alu		1-57 DTHS1	783-828 MIRAGE1		19-56 Coprina_Cc1			1407
DTC37AluF			21-138 CvG			111-182		824
DTC37AluR								813
DTC38Alu		213-282 DTE			183-221 Vingi-1_Tcas			283
DTC39Alu				400-467				1116

				Copia17-VV_I				
DTC40AluF						647-802		870
DTC40AluR								861
DTC41AluF			52-171 Helitron-N3_SP	405-452 RAGYPSY_LTR_MT	588-674 CR1-32_HM	111-206		820
DTC41AluR								860
DTC42AluF					612-669 CR1-44_HM			821
DTC42AluR			739-807 hAT-51_HM					859
DTC43Alu								369
DTC44Alu								662
DTC45Alu			673-759 MuDR-5_VV			332-403	338-403 <i>Corbicula fluminea</i> Msx gen za Msx protein	1124
DTC46Alu						126-224, 255-341	293-339 <i>Corbicula fluminea</i> Msx gen za Msx protein	592
DTC47AluF			282-327 Sola2-2_HM					622
DTC47AluR								810
DTC48Alu						50-241	52-98 <i>Corbicula fluminea</i> Msx gen za Msx protein	545
DTC49Alu						5-84, 127-345		763
DTC50Alu	BIV160 sat					1-323		328
DTC51Alu		1-174 BIV160 sat				1-174		306
DTC52Alu	BIV160 sat					14-720		720
DTC53Dra								249
DTC54Dra								263
DTC55Dra								330
DTC56Dra								263
DTC57Dra							4-200 <i>Escherichia coli</i>	200

							BL21(DE3) replikacijski P protein	
DTC58Dra								263
DTC59Dra								263
DTC60Dra								263
DTC61Dra								266
DTC62Dra								405
DTC63Dra			72-133 Merlin10_SM					263
DTC64Dra								263
DTC65Dra								249
DTC66Dra								263
DTC67Dra								263
DTC68Dra								263
DTC69AluF								917
DTC69AluR								950
DTC70AluF			457-518 DNA-4-4_DR					917
DTC70AluR								910
DTC71Alu								1290
DTC72Alu						212-279		374
DTC73AluF			727-780 hAT-83_HMa					950
DTC73AluR								950
DTC74AluF			722-815 Mariner-17_SM		537-583 CR1-35_HM	515-583		916
DTC74AluR								907
DTC75AluF			148-234 hAT-14_SM			467-557		919
DTC75AluR								907
DTC76Alu								147
DTC77Alu								226

DTC78AluF								919
DTC78AluR								908
DTC79AluF								924
DTC79AluR								805
DTC80Alu								299
DTC81AluF						294-652		922
DTC81AluR			863-945 Polinton-2_NVi					950
DTC82AluF					193-420 SIRE3_TC	144-873		922
DTC82AluR						170-735		906
DTC83Alu						457-492, 551-760	671-777 <i>Corbicula fluminea</i> Msx gen za Msx protein	1133
DTC84Alu			274-433 CVG			255-791	294-1034 DTC84 element	2707
DTC85Alu			74-159 Mariner-N2_SP					968
DTC86Alu								284
DTC1BamF								868
DTC1BamR					179-218 LIN4b_SM			863
DTC2Bam								41
DTC3BamF			24-75 hAT-N3_BF			33-112		869
DTC3BamR			444-490 Helitron-2_DVir					863
DTC4BamF								918
DTC4BamR						588-650		914
DTC5BamF								867
DTC5BamR						587-649		913
DTC6BamF			630-702 ISL2EU-5_HM 738-819 Mariner-N5_DR					868

								Rezultati
DTC6BamR								863
DTC7BamF				337-391 Copia-44_Mad-I				818
DTC7BamR			772-842 MtPH-A6-4-la	149-207 Gypsy-7N_PPa-LTR	267-332 VLINE3_VV			863
DTC8BamF						803-905		918
DTC8BamR								912

Tablica 3.1: Pregled dobivenih genetičkih fragmenata s obzirom na prisutnost različitih repetitivnih elemenata. Brojevi označavaju nukleotidnu poziciju satelitne DNA ili mobilnog elementa unutar klona, cDNA kolona daje pregled odsječaka kojima su pronađeni transkribirani ekvivalenti.

Šesta skupina sekvenci obuhvaća one koje sadrže više od jednog tipa repetitivne DNA, a u njima se mogu pronaći različite kombinacije elemenata prisutnih u svim prethodnim skupinama. Ovdje se nalaze genomske fragmenti: DTC9Alu, DTC17AluF, DTC28Alu, DTC30AluF, DTC32Alu, DTC34Alu, DTC36Alu, DTC38Alu, DTC41AluF, DTC74AluF, DTC84Alu, DTC6BamF i DTC7BamR.

Sedma i posljednja skupina sadrži sekvence koje pokazuju sličnost s retroviralnim elementima (DTC11AluR, DTC25Alu, DTC33Alu i DTC8AluR).

Rezultati u sklopu cDNA kolone Tablice 1 dobiveni pretragom EST kolekcije unutar baze podataka *NCBI GenBank* pokazuju unutar velikog broja klonova područja koja su vrlo vjerojatno transkripcijski aktivna. Ta se područja kod 13 klonova poklapaju s onim istim dijelovima kod kojih je primjećena sličnost s nekim mobilnim elementom, dok je kod nekih utvrđena sličnost i s konkretnim proteinskim produktom, poput reverzne transkriptaze iz transponirajućeg elementa *jockey* (kolona „ostalo“ Tablice 3.1). Tablica 3.2 sumarno prikazuje osnovne tipove mobilnih elemenata s kojima je uočena sličnost unutar genomske fragmenata školjkaša *D. trunculus*, uz izdvojene glavne porodice.

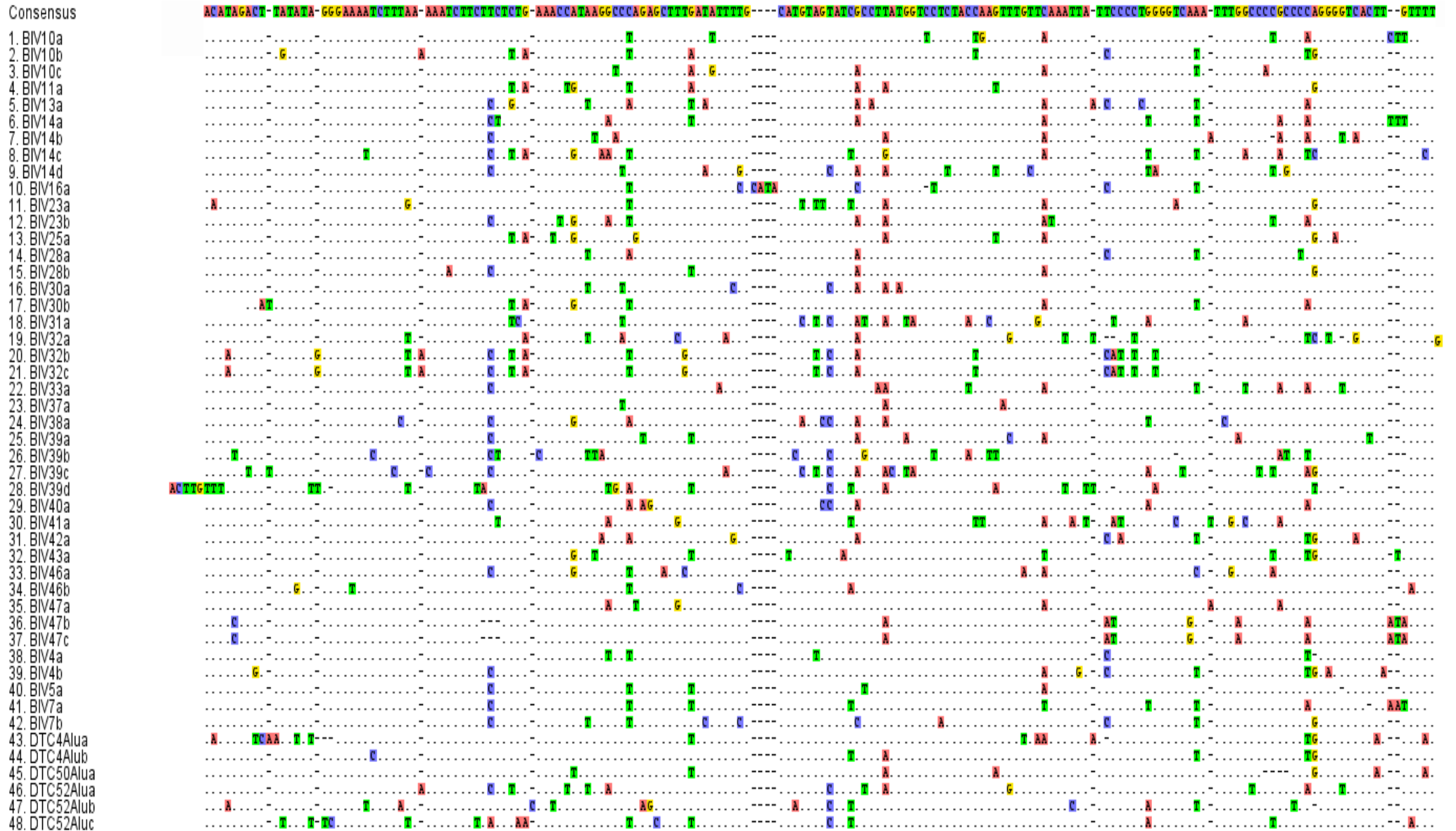
Tip elementa
DNA transpozon (34)
Academ
EnSpm
Harbinger
Helitron
ISL2EU
Mariner/Tc1
Merlin
MuDR
Polinton
Sola
hAT
Retrovirus (4)
ERV1
ERV3
LTR Retrotranspozon (7)
BEL
Copia
Gypsy
Non-LTR Retrotranspozon (16)
CR1
Daphne
L1
L2
NeSL
Penelope
SINE
Vingi
Mobilni elementi (61)

Tablica 3.2: Osnovni tipovi mobilnih elemenata koji su pokazali sličnost s genomskim fragmentima školjkaša *D. trunculus*, uz izdvojene glavne porodice elemenata.

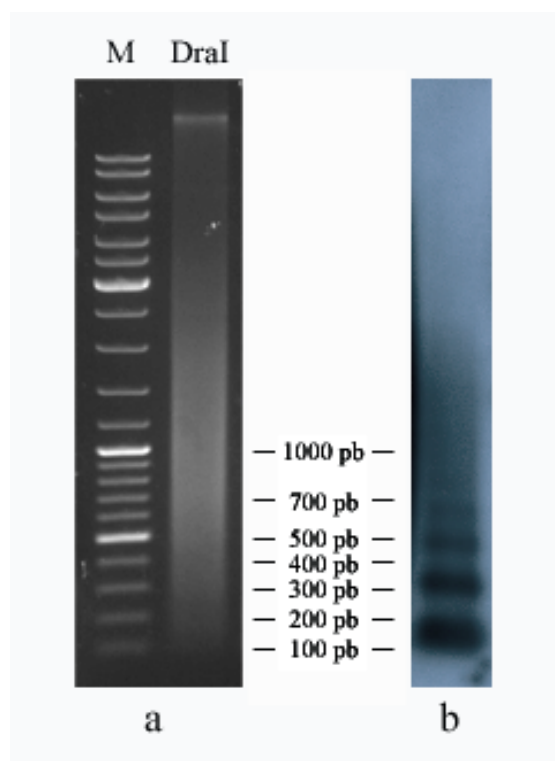
3.1.1. BIV160 porodica satelitnih DNA u školjkašu *Donax trunculus*

Za klon DTC4Alu iz prve skupine te klon DTC17AluF iz šeste skupine, uočeno je da sadrže satelitnu DNA koja pripada BIV160 porodici, staroj 540 milijuna godina i široko rasprostranjenoj među školjkašima (Plohl i sur. 2010). Koristeći monomer-specifičnu hibridizacijsku sondu, nakon analize otisaka kolonija, dobivena su još 3 klona (DTC50Alu, DTC51Alu, DTC52Alu) koja sadrže istu satelitnu DNA. Na temelju postojećih monomera u ovih 5 klonova, konstruirane su početnice specifične za *D. trunculus* varijantu BIV160 satelitne DNA (Dt-BIV160). Dobiveni PCR produkti su klonirani, a multimerna ponavljanja su za potrebe bioinformatičkih analiza izrezana na veličinu monomernih jedinica. Sveukupno je prikupljeno 48 monomera, pri čemu nije uočena razlika među njima s obzirom na različite eksperimentalne pristupe kojima su dobiveni. Svrnjenje svih dobivenih monomera prikazano je na Slici 3.1, uz usporedbu s konsenzus sekvencom. Konsenzus sekvenca je duljine 166 pb, dok se duljine pojedinačnih jedinica ponavljanja kreću u rasponu od 157 pb - 170 pb. AT sastav ovih sljedova iznosi 53,6%, a prosječna varijabilnost među monomerima iznosi 13%.

Osim u obliku kraćeg niza uzastopnih ponavljanja, monomeri BIV160 satelitne DNA nisu pronađeni ni u jednom drugom stupnju organizacije višeg reda (HOR), što je pokazano i genomskom Southern hibridizacijom (Slika 3.2). U svrhu genomske Southern hibridizacije ukupna genomska DNA razgrađena je *DraI* restrikcijom koji ima mjesto prepoznavanja unutar monomera Dt-BIV160 satelitne DNA. Iako se nakon razgradnje s *DraI* i elektroforetskog razvlačenja fragmenata ne može uočiti uobičajeni ljestvičasti profil, karakterističan za uzastopno ponovljenu DNA (Slika 3.2a), nakon hibridizacije sa sondom specifičnom za Dt-BIV160 satelitnu DNA vidljiva je ljestvica čiju osnovnu jedinicu ponavljanja predstavlja fragment duljine približno 160 pb (Slika 3.2b), što odgovara duljini monomera te satelitne DNA.

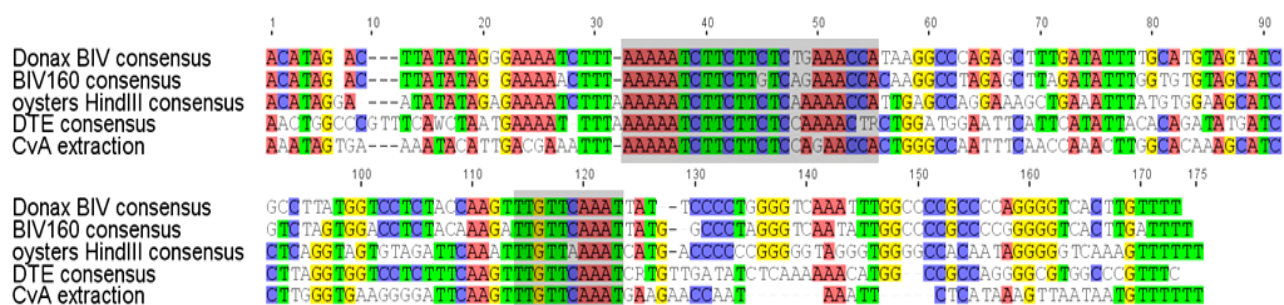


Slika 3.1: Nukleotidni sljedovi jedinica ponavljanja BIV160 satelitne DNA u vrsti *D. trunculus*. Konsenzus sekvenca prikazana je punim nukleotidnim sljedom, dok su kod pojedinih monomera prikazane samo razlike u odnosu na konsenzus.



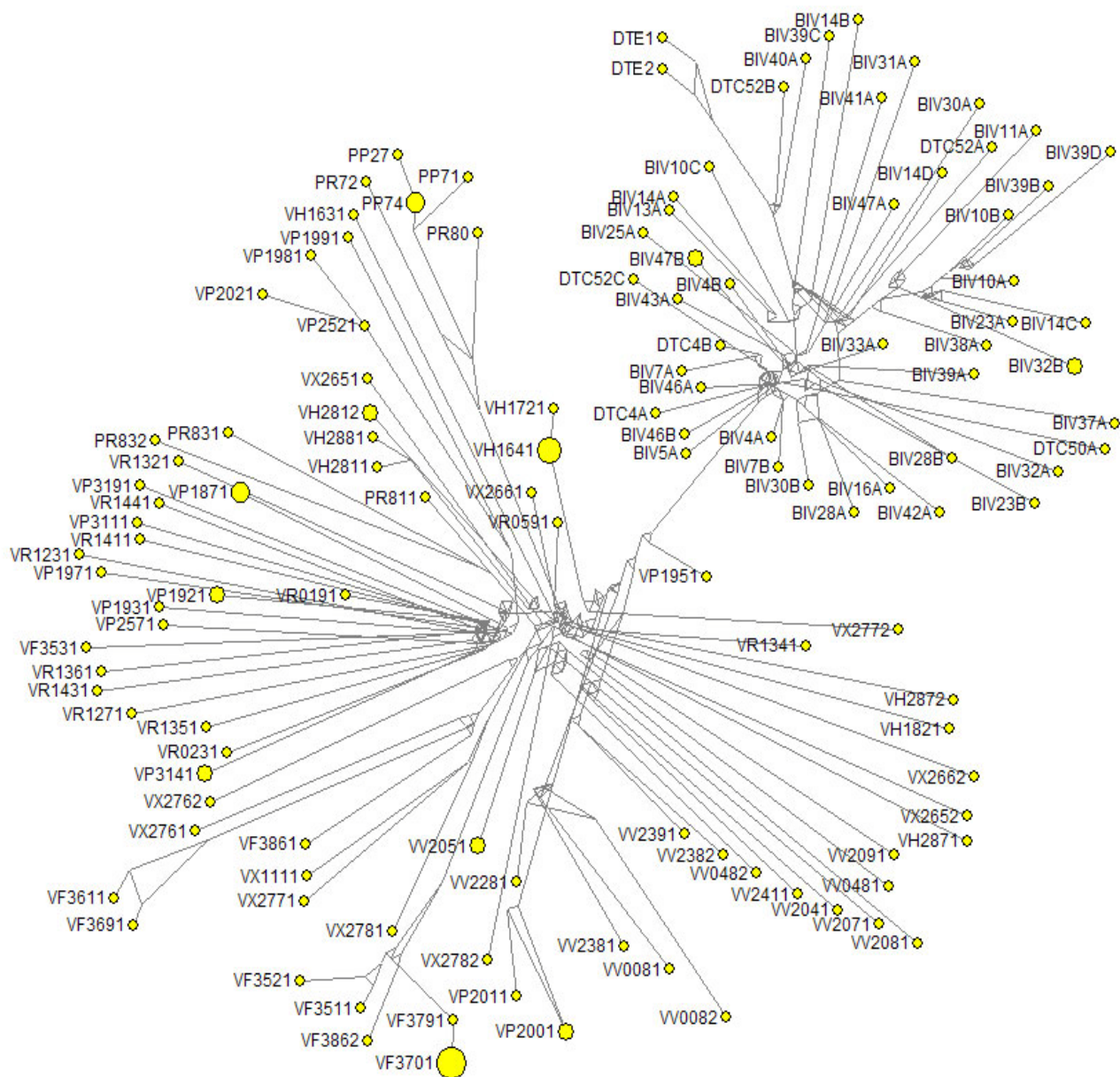
Slika 3.2: Gel elektroforeza genomske DNA razgrađene restriktivnim enzimom *DraI* (a) nakon čega je uslijedila hibridizacija s Dt-BIV160 satelitnom sondom (b).

Stupanj sličnosti između Dt-BIV160 konsenzus sekvence i konsenzus sekvence BIV160 satelitnih DNA prethodno utvrđenih kod nekoliko vrsta školjkaša iznosi 85,6%. S obzirom na postojeću homologiju, Dt-BIV160 satelit je srodan i s mobilnim elementom CvA (*pearl*) iz kamenica (Gaffney i sur. 2003), kao i s nekoliko prethodno opisanih familija satelitnih DNA. Svrnjenje dijela CvA elementa, konsenzus sekvence *D. trunculus* DTE satelitne DNA (Plohl i Cornudella 1996), konsenzus sekvence *HindIII* satelitne DNA nađene kod kamenica (Lopez-Flores i sur. 2004), konsenzus sekvence satelitne DNA BIV160 nađene kod velikog broja školjkaša (Plohl i sur. 2010) te konsenzus sekvence satelitne DNA Dt-BIV160 prikazano je na Slici 3. Konzervirani blokovi prisutni unutar monomera, već ranije utvrđeni kod prve 4 skupine sekvenci, sačuvani su i kod *D. trunculus* varijante BIV160 satelitne DNA.



Slika 3.3: Svrnjnje konsenzus sekvenci satelitnih DNA Dt-BIV160, BIV160, DTE i *HindIII* te dijela CvA mobilnog elementa. Konzervirani blokovi prisutni kod svih 5 skupina sekvenci oznaeni su sivo.

Prilikom usporedbe s varijantama monomera satelitne DNA BIV160 izolirane iz različitih vrsta školjkaša unutar porodice Veneridae, vrsno-specifični Dt-BIV160 monomeri se u filogenetskim analizama odvajaju u zasebnu skupinu, kojoj se priključuje i satelitna DNA DTE (Slika 3.3).

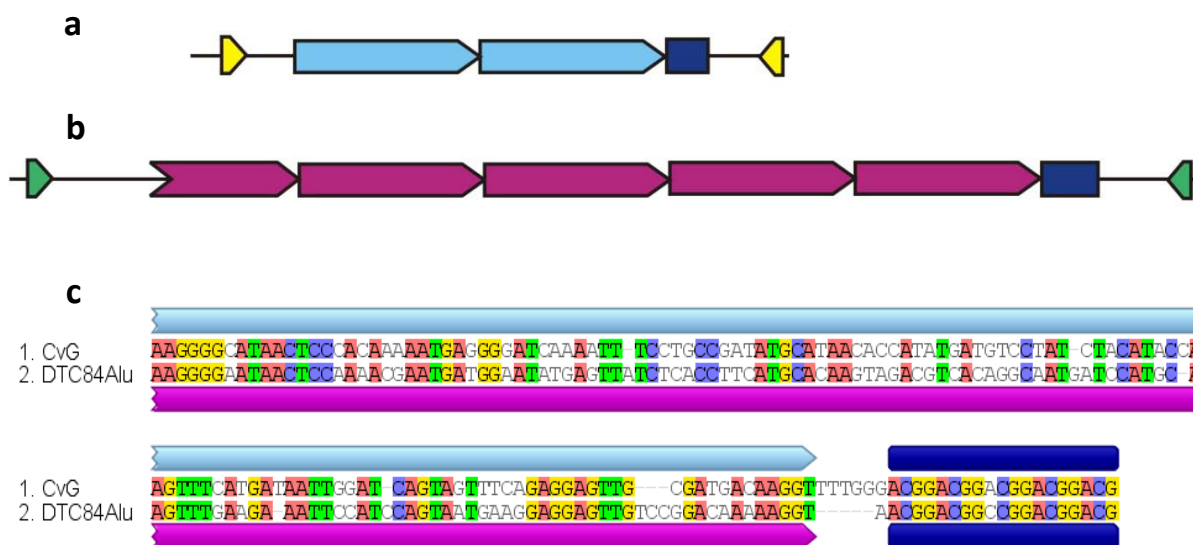


Slika 3.4: Filogenetska mreža monomera BIV160 satelitne DNA. Monomeri koji potječu od vrste *Dosinia exoleta* označeni su slovima VX, od vrste *Ruditapes decussatus* VR i PR, od vrste *Ruditapes philippinarum* VF, od vrste *Venerupis pullastra* VP i PP, od vrste *Venerupis rhomboides* VH, od vrste *Venus verrucosa* VV, od vrste *Donax trunculus* DTC i BIV.

3.1.2. DTC84 element

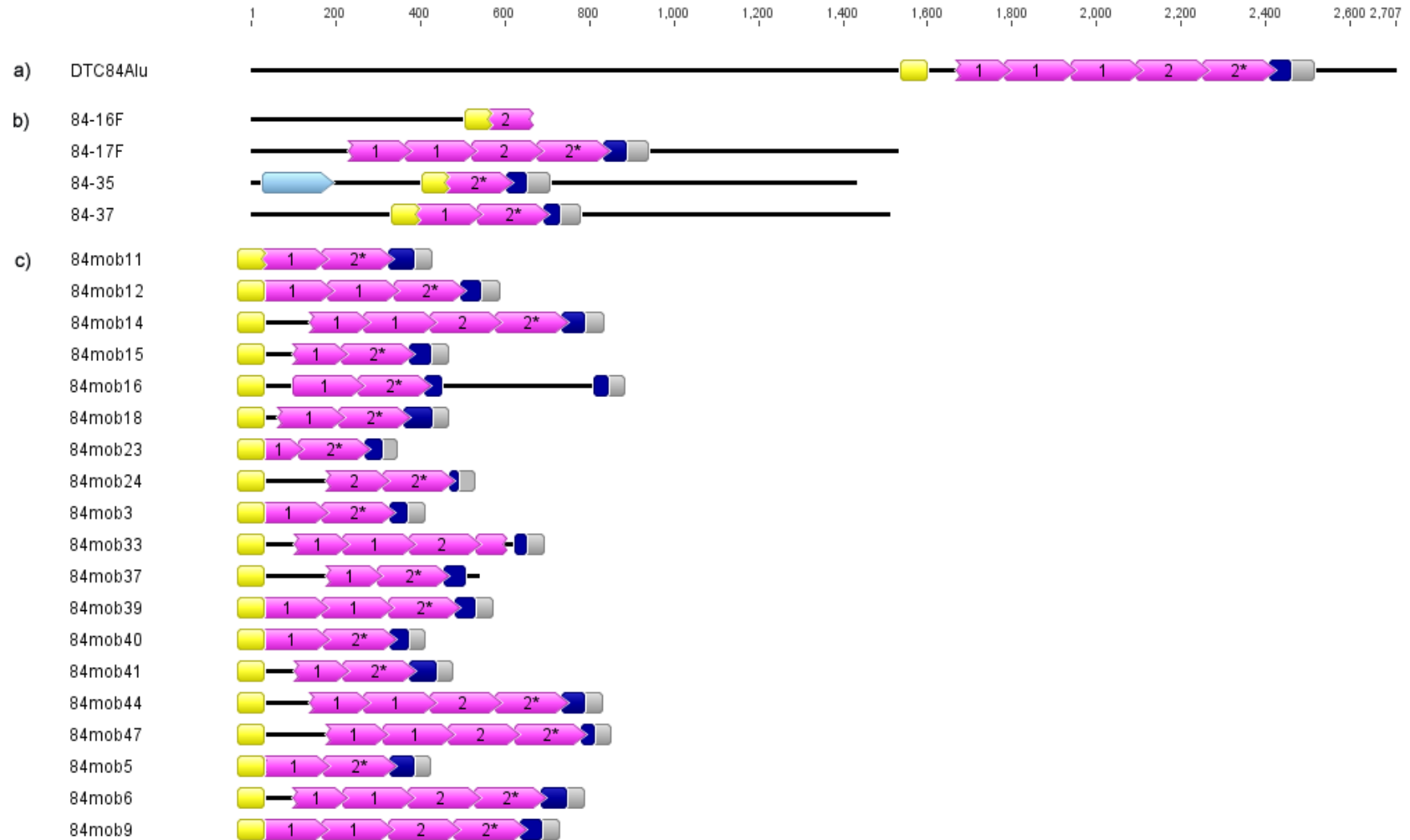
Unutar DTC84Alu klona koji sadrži genomski fragment duljine 2707 pb uočeno je postojanje repetitivne DNA sekvence duljine 741 pb, za koju se činilo da pripada dosad neopisanoj satelitnoj DNA. Detaljnijom analizom te sekvence pokazalo se da se radi o 5 uzastopno ponovljenih jedinica duljine ~160 pb, pri čemu je prva jedinica skraćena na 5' kraju za nekoliko nukleotida, a nakon niza slijedi kratka mikrosatelitna regija. Napravljen je novi ciklus analize otisaka kolonija uz korištenje sonde karakteristične za uzastopna ponavljanja, nakon čega su izdvojena još 4 genomska fragmenta unutar kojih je pronađen isti oblik organizacije kao i kod klona DTC84Alu. Kod svih fragmenata prisutne su jedinice ponavljanja, ili njihovi kratki nizovi, okruženi s konzerviranim blokovima duljine ~50 pb, zbog čega su jedinice dobile naziv „centralna ponavljanja“. Nukleotidni sljedovi koji prethode konzerviranim blokovima, kao i oni koji slijede nakon njih, ne pokazuju nikakvu međusobnu sličnost kod uspoređivanih genomskih fragmenata. U svrhu detaljnije analize, konstruiran je par početnica (DTC84mobF i DTC84mobR), specifičan za konzervirane blokove koji neposredno okružuju centralna ponavljanja ovog elementa. Tim putem je izolirano 19 novih fragmenata koji su pokazivali konzerviranost u nukleotidnom slijedu kao i u strukturnoj organizaciji, a novi element je nazvan DTC84 (prema početnom klonu DTC84Alu). Strukturna organizacija novog elementa DTC84 obuhvaća: konzervirani blok L + (centralna ponavljanja)₁₋₅ + (ACGG)₂₋₁₃ mikrosatelitna ponavljanja + konzervirana regija R. Takva struktura je svojom građom uvelike podsjećala na CvG mobilni element nađen u školjkašu *Crassostrea virginica*, koji je također član ranije spomenute porodice MITE-u sličnih neautonomnih transponirajućih elemenata pod nazivom *pearl* (Gaffney i sur. 2003). Strukturne organizacije elemenata CvG i DTC84 su prikazane na Slici 5a i 5b. Osim u strukturnoj sličnosti postoji određena podudarnost i u nukleotidnom slijedu između elemenata na području centralnih ponavljanja i istovjetne ACGG mikrosatelitne sekvence (Slika 3.5c). Slika 3.6 prikazuje sve DTC84 mobilne elemente, prikupljene različitim eksperimentalnim metodama. Unutar jednog genomskog fragmenta (84-35) izuzev DTC84 elementa, uočeno je prisutnost i ~1 cjelovite monomerne jedinice Dt-BIV160 satelitne DNA, opisane u prethodnom poglavlju.

AA dinukleotid koji se može naći na krajevima DTC84 elementa predstavlja duplikaciju na mjestu (TSD) i ukazuje na to da je vrlo vjerojatno transpozicija mehanizam koji je odgovoran za širenje tih elemenata. Tome u prilog ide i postojanje subterminalnih invertiranih ponavljanja duljine 11 pb koja su smještena unutar konzerviranih blokova L i R. Invertirano ponavljanje je smješteno subterminalno u bloku L i terminalno u bloku R, što je u skladu s njihovom pozicijom kod CvG elementa (Slika 3.5a).



Slika 3.5: Organizacijska struktura CvG (a) i DTC84 (b) elemenata. Žuto i zeleno označava invertirna ponavljanja, plavo i ružičasto centralna ponavljanja, a tamno plavo mikrosatelitne sekvence. Svrnjenje nukleotidnih sljedova dijela CvG elementa i dijela DTC84 elementa. (c)

Prema onome što slijedi nakon konzerviranog bloka L elementi se mogu grupirati u 3 skupine. Kod prve skupine centralna ponavljanja započinju nastavno na blok L, dok preostale dvije skupine sadrže nukleotidne sljedove duljine ~80 pb ili ~160 pb prije početka samih centralnih ponavljanja, pri čemu im je prvih 16 nukleotida zajedničko (Slika 3.7). Na drugom kraju elementa, u nastavku mikrosatelitnog niza direktno slijedi konzervirani blok R. Iznimke su uočljive samo kod 84mob16 i 84mob33, kod kojih je uobičajena struktura narušena insercijama i delecijama.



Slika 3.6: DTC84 mobilni elementi prikupljeni različitim eksperimentalnim metodama. a) dobiven nakon hibridizacije sa kompletnom genomskom DNA kao sondom. b) dobiveni nakon analize otisaka kolonija sa sondom specifičnom za centralna ponavljanja c) dobiveni umnažanjem sa DTC84mobF i DTC84mobR početnicama. Blok L je označen žuto, blok R sivo, mikrosatelitna regija tamno plavo, centralna ponavljanja ružičasto. Brojke 1 i 2 označavaju tip kojemu centralna ponavljanja pripadaju, 2* označava posebnu podvrstu tipa 2 (vidi tekst). Svjetlo plavo označeni dio klona 84-35 pripada satelitnoj DNA Dt-BIV160 (vidi tekst).



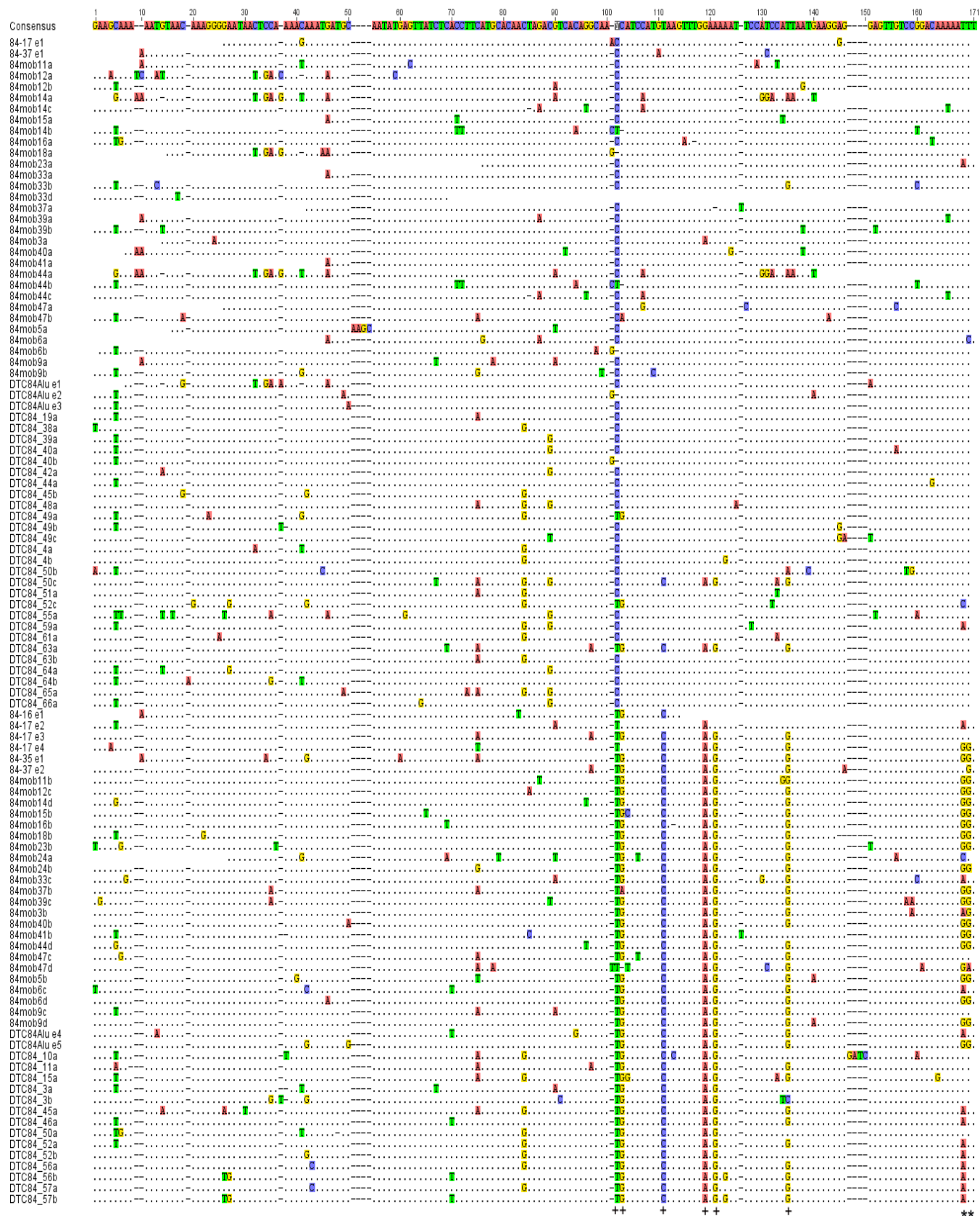
Slika 3.7: Nukleotidni sljedovi koji prethode centralnim ponavljanjima. Blok L je označen žuto, mjesto za DTC84mobF početnicu crveno, invertno ponavljanje zeleno.

U svrhu daljnje analize DTC84 elementa u genomu školjkaša *D. trunculus* konstruiran je set početnica (DTC84AluSatF i DTC84AluSatR), specifičan za umnažanje centralnih ponavljanja. Dobiveni PCR produkti su pokazali ljestvičasti raspored elektroforetskih pruga, karakterističan za repetitivnu DNA, pri čemu nisu uočene elektroforetske pruge koje bi odgovarale većem broju ponavljanja od onih dosad pronađenih unutar elemenata. Dobiveni PCR produkti su klonirani, a iz multimernih nizova je izdvojeno 41 novih jedinica ponavljanja, koja su korištena u daljnjim analizama zajedno s 65 centralnih ponavljanja izdvojenih iz kloniranih elemenata. Stupanj homologije između uzastopnih ponavljanja dobivenim PCR-om iznosi 94%, a između centralnih ponavljanja izdvojenih iz elemenata 92,6%. Sravnjenje svih jediničnih ponavljanja, prikupljenih različitim metodama, prikazano je na Slici 9. Konsenzus duljina jedinice ponavljanja iznosi 156 pb. Kod jedinica ponavljanja izdvojenih iz DTC84 elemenata, prvo ponavljanje je redovito skraćeno na 5' kraju, i to u rasponu od 4-67 nukleotida. Samo su 2 početna centralna ponavljanja (od njih 24), zadržala svoju cjelovitost (84mob12a i 84mob16a). Skraćene varijante započinju na 11 različitih nukleotidnih pozicija unutar sekvence centralnog ponavljanja, pri čemu 7 jedinica započinje na 8. nukleotidnoj poziciji, četiri započinju na 46., dok se 1-2 ponavljanja mogu naći na svim ostalim startnim pozicijama. Također je uočeno da su one varijante koje započinju na poziciji 8 direktno vezane na konzervirani blok L, dok onima koji započinju na poziciji 46 prethodi kraća varijanta prijelazne sekvence. Sva centralna ponavljanja pri svojem 3' kraju imaju 12-nukleotidnu palindromsku sekvencu TTGTCCGGACAA. Ona pokazuje visoku očuvanost kod većine centralnih ponavljanja, a promijenjene varijante se mogu naći kod 16 od 106 slučajeva.

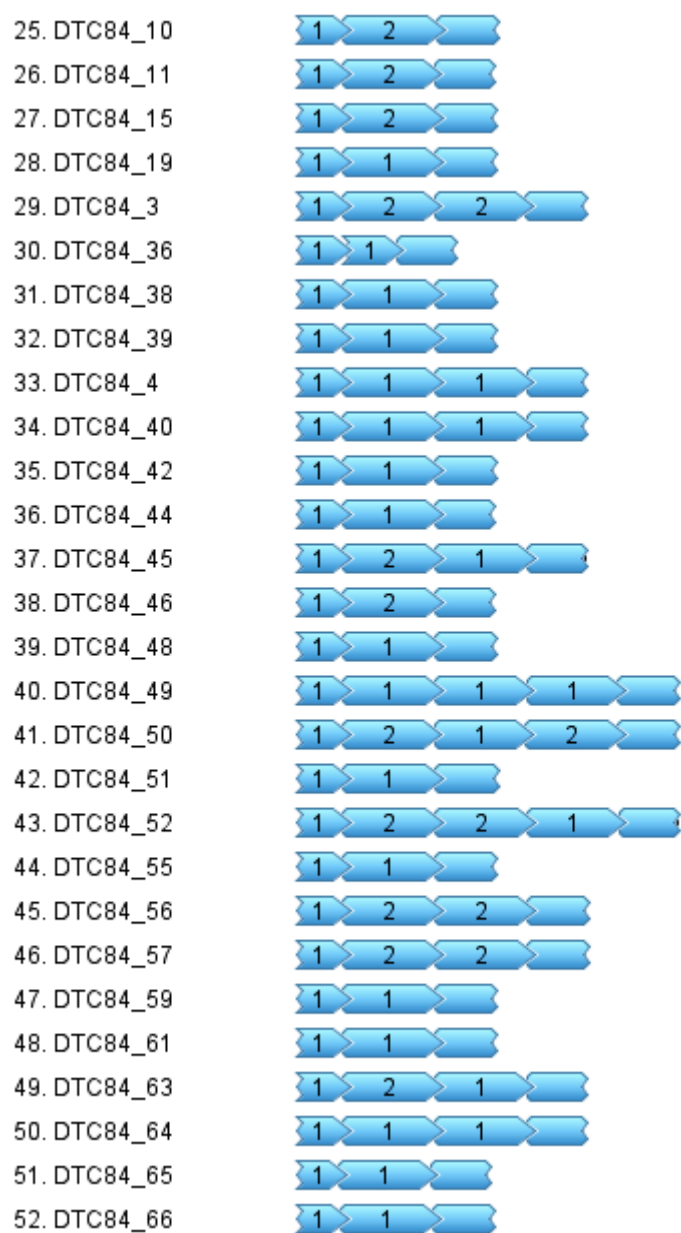
Prema razlikama u nukleotidnoj sekvenci centralna ponavljanja se mogu svrstati u 2 grupe. Postoji 6 nukleotidnih pozicija koje bi se mogle uzeti kao dijagnostičke za pripadnost grupi 1 ili 2, dok su ostale razlike posljedice nasumično nakupljenih mutacija. Kod grupe 1 to su CC-T-G-A-T nukleotidi, a kod grupe 2 TG-C-A-G-G (Slika 3.9). Raspored varijanti 1 i 2 unutar mobilnih elemenata je vidljiv na Slici 6. Može se zamijetiti da se ponavljanja prve grupe uglavnom nalaze na početku niza, dok se grupa 2 uvijek nalazi uz mikrosatelitni dio elementa. U nekoliko slučajeva niz čine samo varijante grupe 2, kao kod 84-35, 84mob24, i kod parcijalnog elementa u genomskom fragmentu 84-16F. U skladu s primjećanim trendom da je prvo ponavljanje u nizu obično skraćeno, u ovim slučajevima to je monomer grupe 2. Pri tomu, monomeri uz mikrosatelitnu regiju, čine podvarijatu 2*, koja se od ostalih razlikuje

prema dva dijagnostička nukleotida pozicionirana na samom kraju, na prijelazu u mikrosatelitnu sekvencu (Slika 3.6, Slika 3.9).

Postojanje grupa 1 i 2 je prisutno i kada su varijante dobivene početnicama karakterističnim za centralna ponavljanja, pri čemu njihov raspored ne poštuje iste organizacijske principe kao jedinice prisutne unutar mobilnih elemenata (Slika 3.10). Zbog drugačijeg eksperimentalnog pristupa kojim su dobiveni, postoji mogućnost da multimeri s izmjeničnim rasporedom varijanti potječu iz nizova koji bi se potencijalno nalazili u nekom organizacijskom obliku izvan DTC84 elemenata, u zasad neutvrđenom genomskom okolišu.

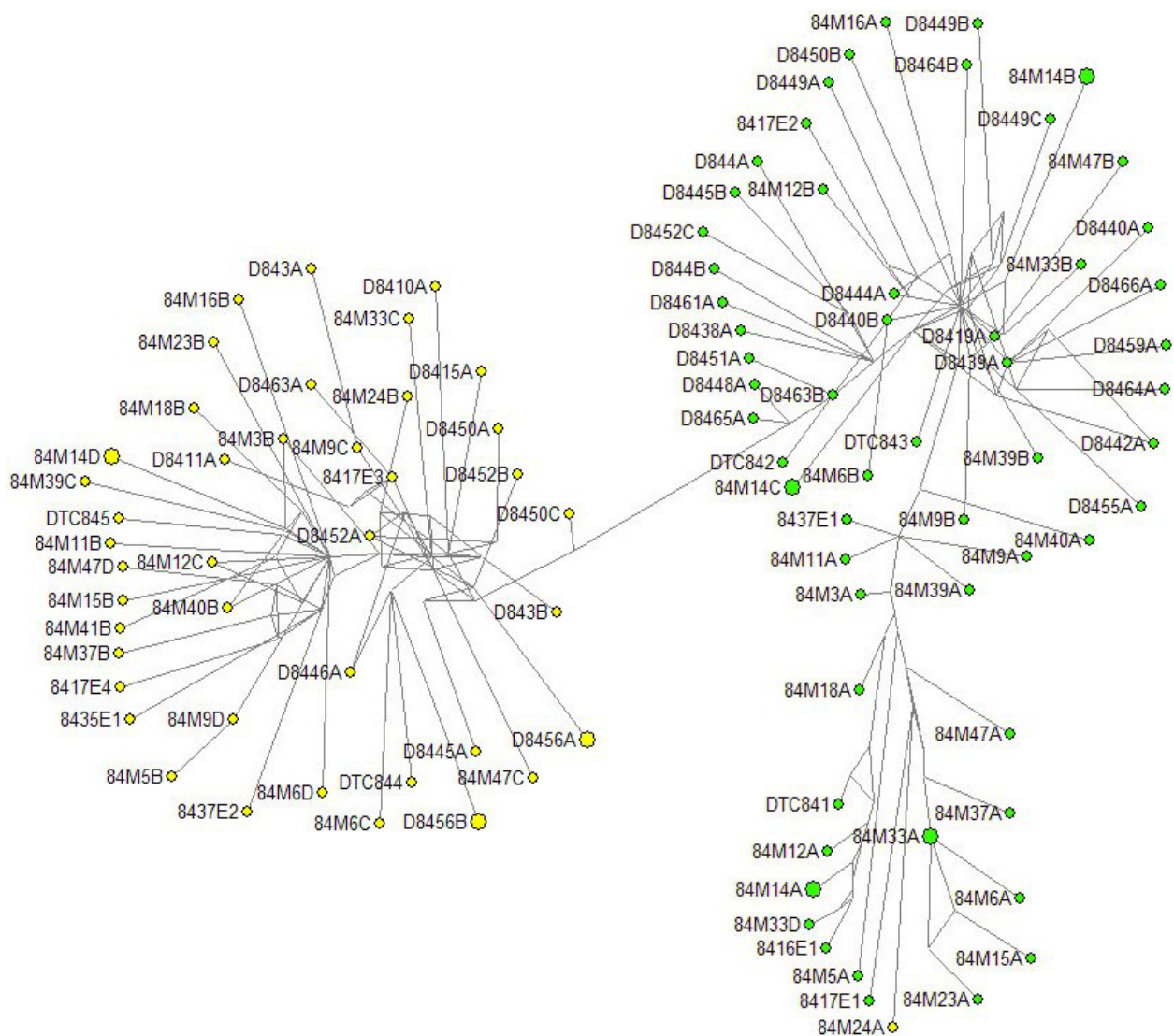


Slika 3.9: Svrstavanje svih jediničnih ponavljanja, prikupljenih različitim metodama. Konsenzus sekvenca prikazana je punim nukleotidnim slijedom, dok su kod ostalih prikazane samo razlike u odnosu na konsenzus. Šest dijagnostičkih pozicija, bitnih za svrstavanje u grupu 1 ili 2, nose oznaku + na dnu slike. Dva dijagnostička nukleotida specifična za podvarijantu 2* označena su zvjezdicom.



Slika 3.10: Uzastopna ponavljanja dobivena nakon umnažanja s DTC84AluSatF i DTC84AluSatR početnicama te raspored varijanti unutar njih.

Filogenetske analize potvrđuju svrstavanje jedinica ponavljanja u 2 osnovne grupe (Slika 3.11). Lijeva grana predstavlja jedinice ponavljanja grupe 2, a desna grupe 1. Unutar grupe 1 izdvaja se podgrupa centralnih ponavljanja koja su skraćena na 5' kraju, pri čemu je jedini izuzetak 5' skraćeni 84M24A, koji s obzirom na dijagnostičke nukleotide pripada grupi 2.



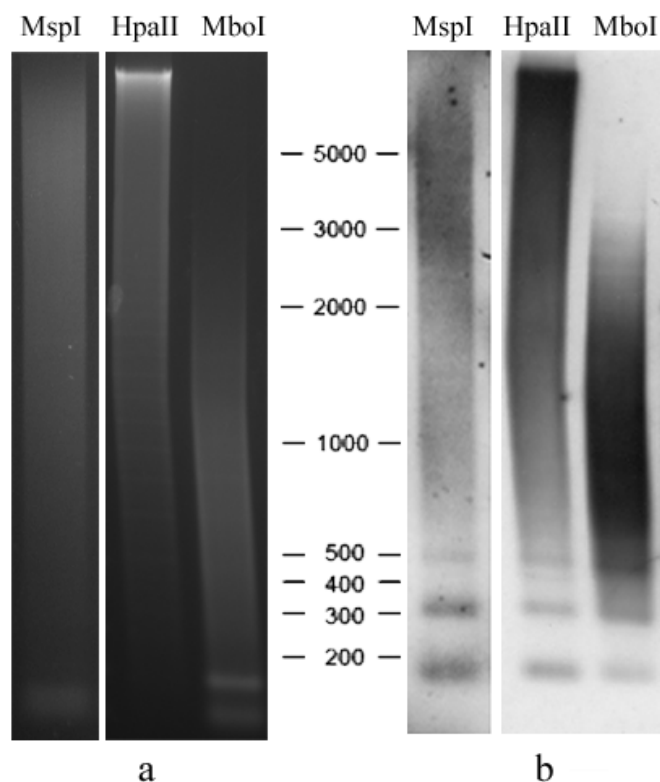
Slika 3.11: Filogenetska mreža jedinica izdvojenih iz uzastopnih ponavljanja dobivenih PCR-om i centralnih ponavljanja izdvojenih iz DTC84 elemenata. Prema svrstavanju jedinica ponavljanja na Slici 9, one koje pripadaju grupi 1 su označene zeleno, a one koje pripadaju grupi 2 žuto. Uočena razlika je samo u 84M24A (objašnjenje u tekstu).

3.1.2.1. Organizacija i udio DTC84 ponavljanja u genomu školjkaša *Donax trunculus*

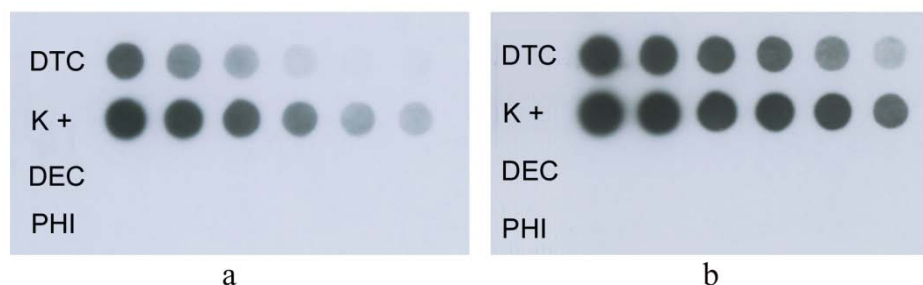
U svrhu dobivanja boljeg uvida u organizaciju DTC84 ponavljanja u genomu, primjenjena je genomski Southern hibridizacija. DNA školjkaša je razgrađena s 3 restrikcijska enzima. *MspI* i *HpaII* su endonukleaze koja režu jednom unutar DTC84 jedinice ponavljanja dok endonukleaza *MboI* preferencijalno reže unutar ponavljanja tipa 2. Da bi se ustanovilo koji od restrikcijskih fragmenata sa Slike 12a zaista pripadaju DTC84 ponavljanjima i uzevši u obzir činjenicu da mjesta prepoznavanja tih enzima postoje kod nekoliko od ranije poznatih satelitnih DNA vrste *D. trunculus*, napravljena je hibridizacija po Southernu. U hibridizaciji je korištena smjesa dimernih DTC84 ponavljanja prethodno obilježenih digoksigeninom. Pokazano je da sve restrikcijske razgradnje pokazuju pravilne ljestvice na osnovi fragmenta duljine ~ 160 pb (Slika 3.12b). Dobiveni profil je u skladu sa organizacijom centralnih ponavljanja u kratke nizove u sklopu DTC84 elemenata, dok postojanje razmaza fragmenata DNA različite veličine, posebno na području većih molekulskih masa, je moguća posljedica asocijacije centralnih ponavljanja ili njihovih fragmenata s drugim genomskim sljedovima. Također, valja spomenuti da su enzimi *MspI* i *HpaII* izoshizomeri, pri čemu *MspI* nije osjetljiv na metilaciju CpG mjesta dok ista blokira *HpaII* aktivnost. Opažene razlike u restrikcijskim i hibridizacijskim profilima (Slika 3.12b) su u skladu s ranije primjećenom pojavom metiliranosti DNA ovog školjkaša (Petrović et al. 2009).

Udio sekvenci karakterističnih za centralna ponavljanja u ukupnoj genomskoj DNA određena je pomoću kapljične (eng. *dot blot*) analize (Slika 3.13). Analiza je provedena u različitim temperaturnim uvjetima koji omogućuju sparivanje nukleotidnih sljedova različitih stupnjeva homologije. Viša temperatura podrazumijeva i strože uvjete, odnosno sparivanje sonde i DNA odsječaka visokog postotka homologije. Nakon denzitometriranja hibridizacijskih signala uspoređeni su signali razrjeđenja genomske DNA vrste *D. trunculus* i pozitivne kontrole poznate koncentracije. Pri 68°C, u strožim hibridizacijskim uvjetima (Slika 3.13a), ustanovljeno je da uzastopni sljedovi sveukupno čine oko 1% genoma, tj. prisutni su u $8,9 \times 10^4$ kopija, uzevši u obzir veličinu genoma utvrđenu od strane Hinegardnera (1974). Njihov doprinos genomu pokazuje približno dvostruku vrijednost u manje restriktivnim uvjetima hibridizacije pri 60°C (Slika 3.13b), što ukazuje na raširenost njima srodnih sekvenci u genomu školjkaša. Osim genomske DNA vrste *D. trunculus*, u analizu su uvrštene genomske DNA još 2 vrste školjkaša, *Ruditapes decussatus* i *Ruditapes philippinarum*, kod kojih nije uočena prisutnost DNA sljedova od interesa čak ni u blažim uvjetima hibridizacije (Slika 3.13 a i b).

S obzirom na hibridizacijske profile uočene nakon genomske Southern hibridizacije, teško je procijeniti koliki se udio tih sekvenci nalazi u sklopu centralnih ponavljanja u odnosu na one koji se nalaze u asocijaciji s drugim genomskim sljedovima. U svakom slučaju, moguće je zaključiti da centralna ponavljanja u sklopu DTC84 elemenata sačinjavaju <1% genoma.



Slika 3.12: Gel elektroforeza genomske DNA razgrađene restriktivnim enzimima MspI, HpaII, MboI (a) nakon čega je uslijedila hibridizacija sa sondom specifičnom za centralna ponavljanja elementa DTC84 (b).

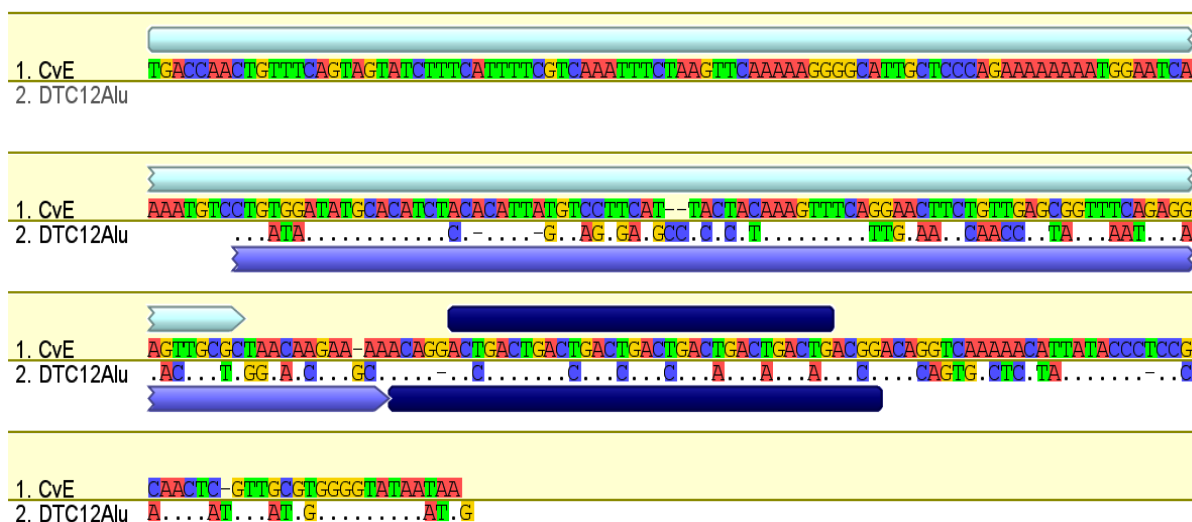


Slika 3.13 : Određivanje udjela sekvenci karakterističnih za centralna ponavljanja elementa DTC84 u ukupnoj genomskoj DNA vrsta *D. trunculus* (DTC), *R. decussatus* (DEC) i *R. philippinarum* (PHI) pomoću kapljичne analize, pri različitim uvjetima hibridizacije : 68°C (a) i 60°C (b).

3.1.3. Ostale sličnosti s elementima porodice *pearl*

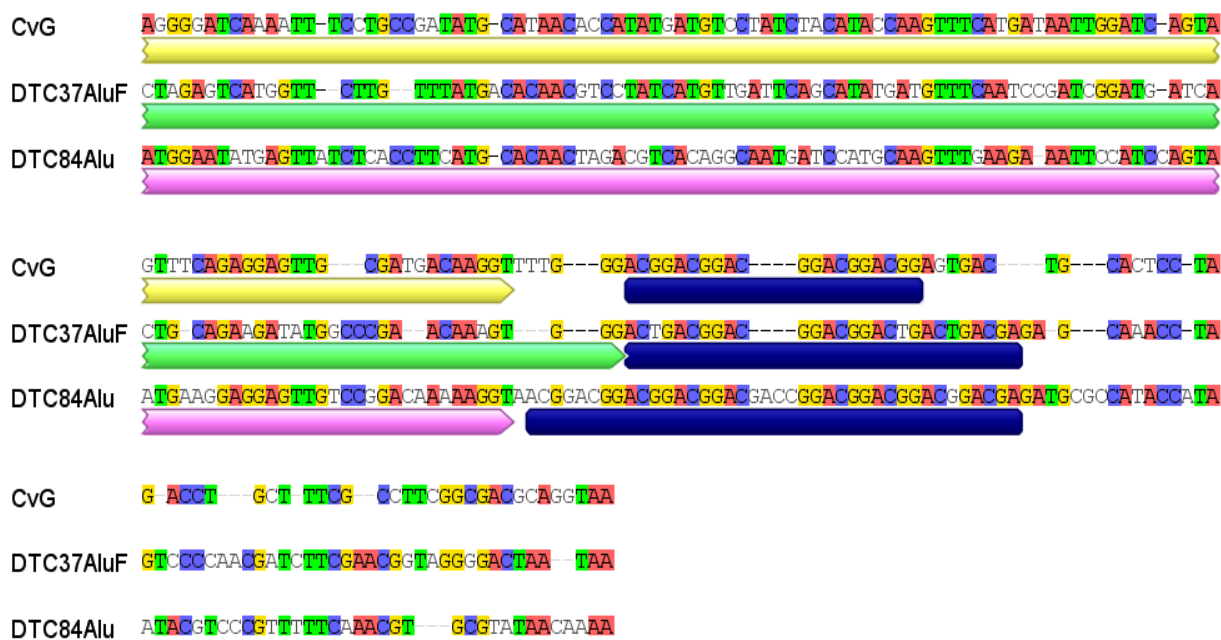
U tablici 3.1 uočeno je postojanje još 2 genomska fragmenta koji pokazuju sličnost s elementima porodice *pearl*, DTC12Alu i DTC37AluF.

Genomski klon DTC12Alu sadrži insert duljine 190 pb i pokazuje 63.2% sličnosti s elementom CvE porodice *pearl* (Gaffney i sur. 2003; Slika 3.14). Područje preklapanja obuhvaća dio centralnog ponavljanja, mikrosatelitnu regiju te regiju koja slijedi nakon nje. Dok se mikrosatelitna ponavljanja elementa CvE sastoje od motiva ACTG, kod klona DTC12Alu mogu se naći malo promijenjene varijante, 5 puta ponovljena varijanta ACCG i 4 puta ponovljena varijanta ACAG.



Slika 3.14: Svrnjenje nukleotidnih sljedova dijela elementa CvE i klona DTC12Alu. Centralna ponavljanja označena su svjetlo plavo i ljubičasto, mikrosatelitne regije tamno plavo.

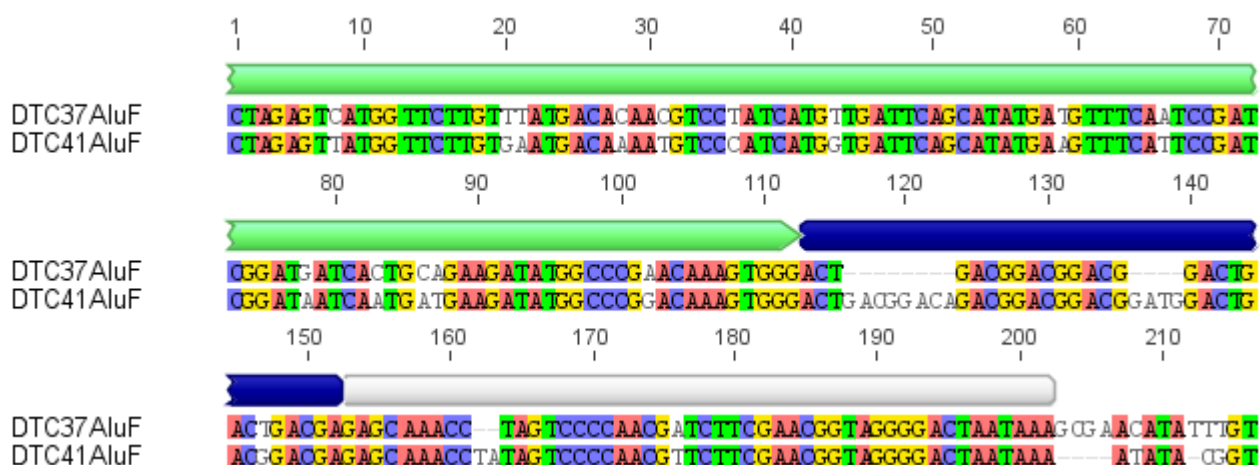
Kod sljedećeg klona, DTC37Alu, početnih 140 nukleotida njegovog inserta pokazuje ~61% sličnosti s centralnim ponavljanjima elemenata DTC84 i CvG te njihovim pripadajućim mikrosatelitnim regijama (Slika 3.15). To ga stavlja u poziciju jedne od divergentnih varijanti iz porodice DTC84-sličnih elemenata i djelomično objašnjava značajne razlike u udjelu tih sljedova u genomu školjkaša pri manje restriktivnim uvjetima hibridizacije (Slika 3.13 a i b).



Slika 3.15: Svrnjenje nukleotidnih sljedova dijelova elementa CvG, klona DTC37AluF i elementa DTC84. Centralna ponavljanja označena su žuto, zeleno i ružičasto, mikrosatelitne regije tamno plavo.

Mikrosatelitna regija unutar klona DTC37Alu pokazuje heterogenost jer sadrži 7 ponavljanja sastavljena od 3 različita motiva. Pritom se tetranukleotid ACGG (koji je prisutan kod elemenata DTC84 i CvG) nalazi u centralnom dijelu mikrosatelitne regije i ponavlja se 3 puta, a okružen je motivima ACTG i ACGA.

Local blast pretraga pronalazi još jedan nukleotidni slijed visokosrodan elementu pronađenom u sklopu fragmenta DTC37AluF. Element se nalazi u početnom dijelu genomske fragmenta DTC41AluF i također nije cjelovit, odmah počinje unutar sekvence centralnog ponavljanja, na istoj poziciji kao i kod DTC37AluF (Slika 3.16).



Slika 3.16: Svrnjenje nukleotidnih sljedova necjelovitih elemenata pronađenih u sklopu genomskih fragmenata DTC37AluF i DTC41AluF. Dio centralnih ponavljanja je označen zeleno, mikrosatelitne regije tamno plavo, konzervirana regija sivo.

Na području 110 pb samog centralnog ponavljanja nukleotidni sljedovi pokazuju međusobnu sličnost od 87%, nakon čega slijedi mikrosatelitna regija koja se razlikuje po duljini (7 tetranukleotidnih ponavljanja kod DTC37AluF i 10 tetranukleotidnih ponavljanja kod DTC41AluF). Iza mikrosatelitne regije slijedi dobro sačuvani dio od 50 pb (94% homologije), a koji se razlikuje od regija koje se na tom mjestu mogu naći kod srodnih elemenata DTC84 i CvG.

3.1.4. Sličnost satelitnih DNA školjkaša *Donax trunculus* i različitih skupina mobilnih elemenata

Sličnost *HindIII*, DTE i BIV160 porodica satelitnih DNA s CvA mobilnim elementom je poznata od ranije (Gaffney i sur. 2003, Plohl i sur. 2010). U ovom radu im je priključena i satelitna DNA Dt-BIV160 (Slika 3.3), no to nisu jedine porodice satelitnih DNA u školjkašu *D. trunculus* koje pokazuju sličnost s mobilnim elementima. Pretragom baze podataka *Rebase* uočeno je da je to slučaj s još 2 od ranije poznate porodice.

Jedna od njih je DTSH1 (Plohl i Cornudella 1997). Fragmenti DTSH1 monomera su također pronađeni i u 2 genomska fragmenta analizirana u sklopu ovog istraživanja (DTC34Alu i

DTC36Alu, Tablica 3.1), a pokazuju sličnost s početnim dijelom non-LTR retrotranspozona Coprina_Cc1 koji je uglavnom vezan za telomerne sekvence te je često skraćen na 5' kraju zbog ubacivanja telomernih ponavljanja (Gladyshev i Arkhipova 2007).

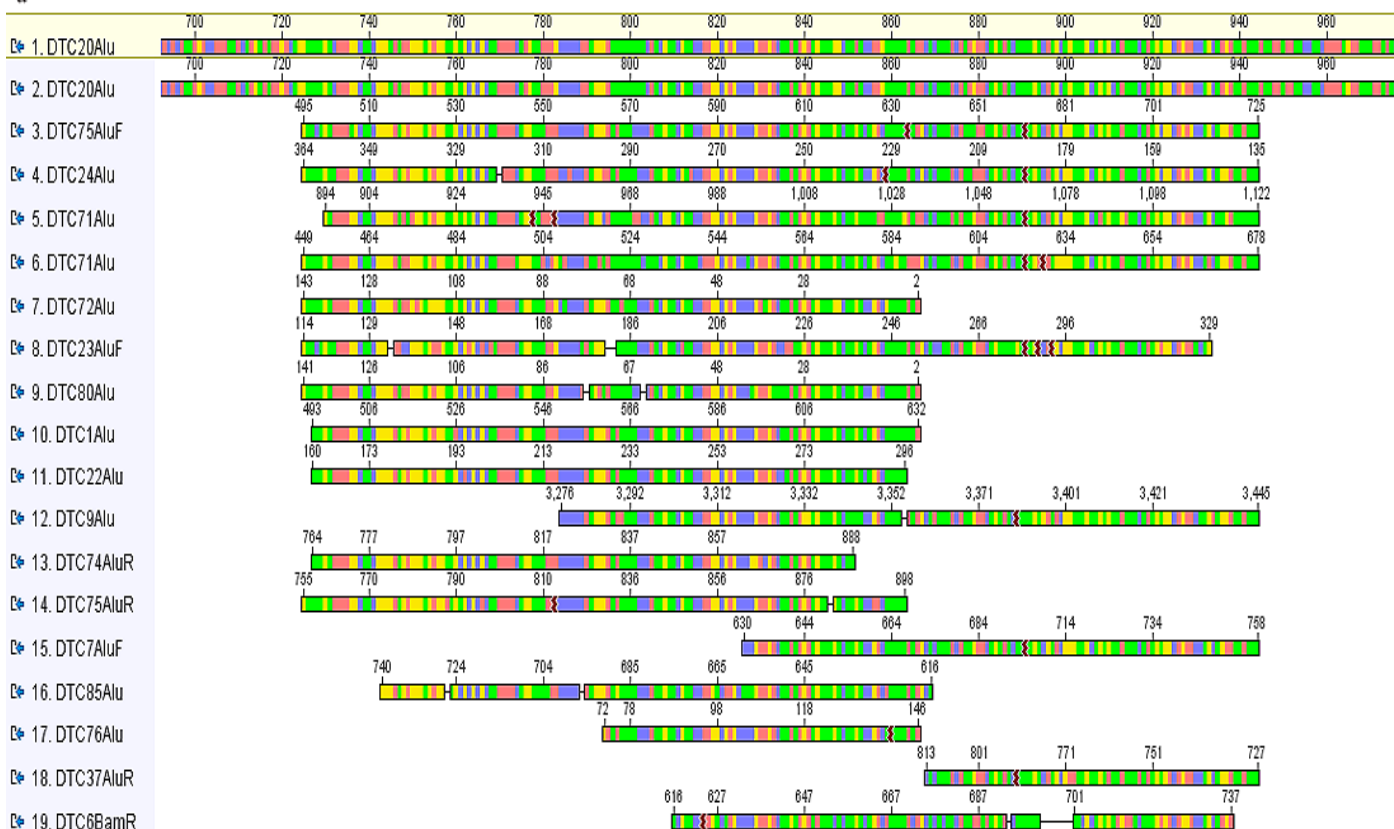
Druga porodica je DTHS2 (Plohl i Cornudella 1997), koja cijelom svojom dužinom monomera pokazuje relativno niski stupanj homologije (62%) s dijelom LTR retrotranspozona Gypsy-144_AA-I (Nene i sur. 2007).

3.1.5. Grupiranja sekvenci

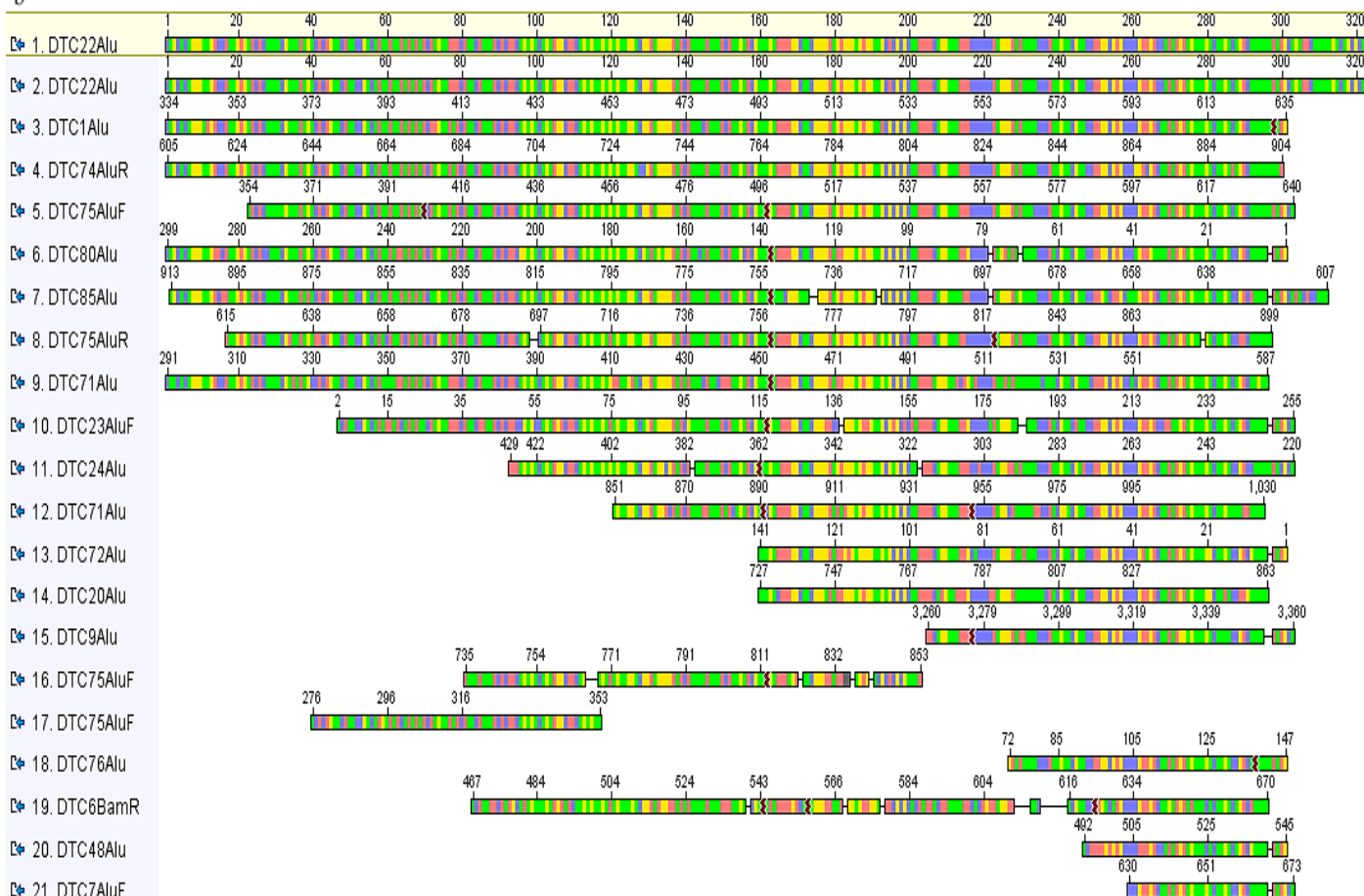
Local blast pretraga koja pronalazi regije sličnosti između nukleotidnih sljedova provedena je unutar cjelokupne baze genomskih fragmenata dobivenih u ovom istraživanju (Tablica 3.1), a kao rezultat uočena su brojna grupiranja sekvenci. Neka od uočenih grupiranja prikazana su na Slici 17. Takva grupiranja ukazuju na pojavnost određenih nukleotidnih sljedova unutar većeg broja genomskih klonova, a s obzirom da se pojavljuju u različitim genomskim okruženjima, možemo zaključiti da se radi o nekom obliku raspršeno ponovljenih sekvenci. Kao što je vidljivo na Slici 18, ponekad se unutar istog klona, u bliskom genomskom okruženju, može pojaviti više raspršenih ponavljanja koja mogu biti istog ili različitog tipa (objašnjenje kasnije u tekstu).

Detaljnijom analizom grupiranja, kod 16 klonova uočena je prisutnost većeg broja ponovljenih genomskih fragmenata koji su izdvojeni kao reprezentativni i podvrgnuti daljnjim analizama. Slika 3.19a pokazuje izdvojene dijelove tih fragmenata. Brojevima je označeno područje dijela inserta na kojem se uočena ponavljanja nalaze, a Slika 3.19b pokazuje dio Tablice 3.1 s osnovnim informacijama o tim klonovima. Iz slike je vidljivo da se samo za jedan izdvojeni segment klona DTC26Alu može naći podudarnost u tablici i to s dijelom gena *Msx* školjkaša *Corbicula fluminea*. Dodatnom pretragom baza podataka *NCBI GenBank* i *Rebase* nije nađena nikakva druga podudarnost s bilo kojom od izdvojenih sekvenci DNA. To ukazuje na činjenicu da se radi o novim i dosad nepoznatim repetitivnim sekvencama i/ili mobilnim elementima pa su neka od grupiranja detaljnije istražena.

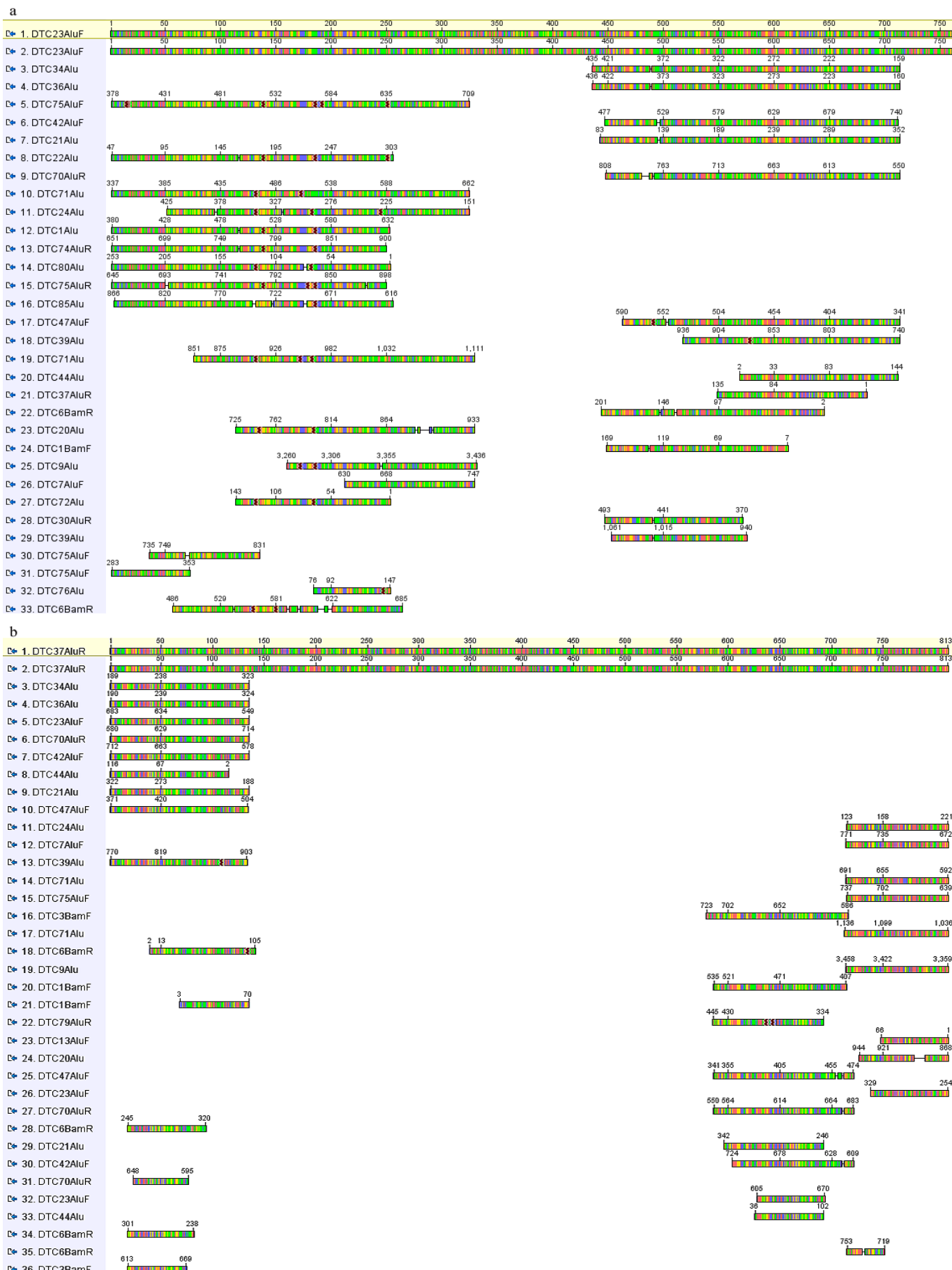
a



b



Slika 3.17: Rezultat *Local blast* analize u kojoj je zbirka genomskih fragmenata izoliranih tijekom ovog istraživanja pretražena klonovima DTC20Alu (a) i DTC22Alu (b)



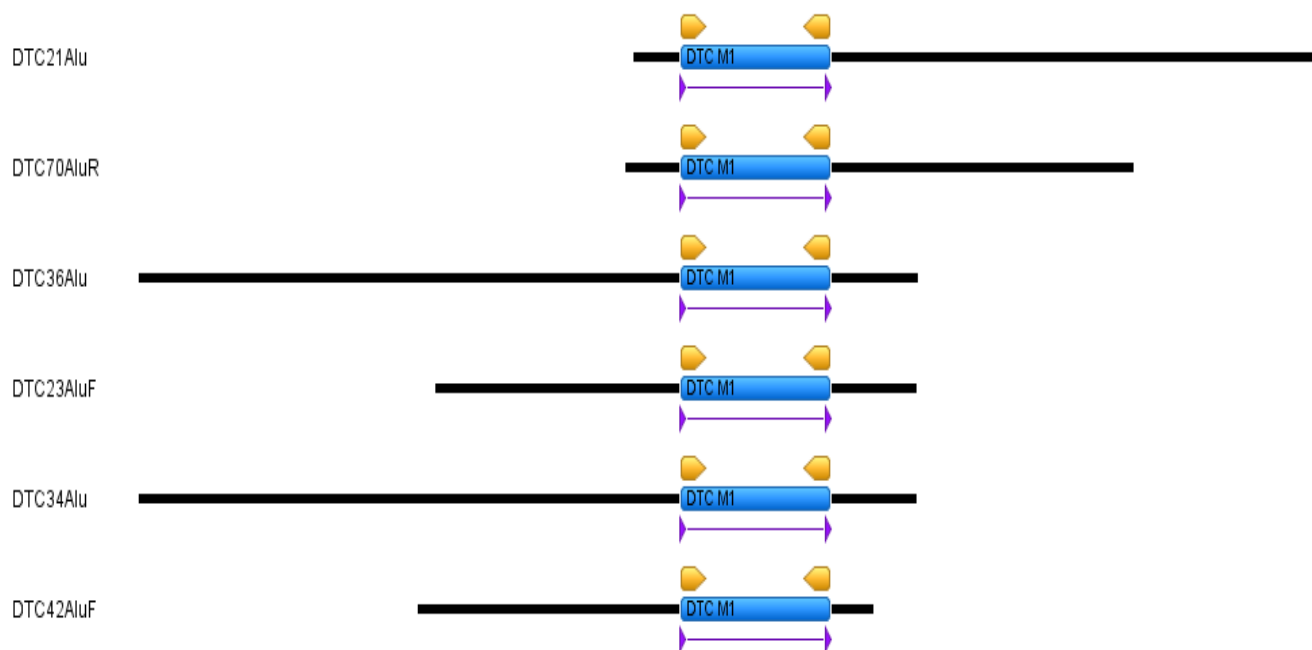
Slika 3.18: Rezultat *Local blast* analize u kojoj je zbirka genomskih fragmenata izoliranih tijekom ovog istraživanja pretražena klonovima DTC23AluF (a) i DTC37AluR (b)

3.1.6. DTC M1

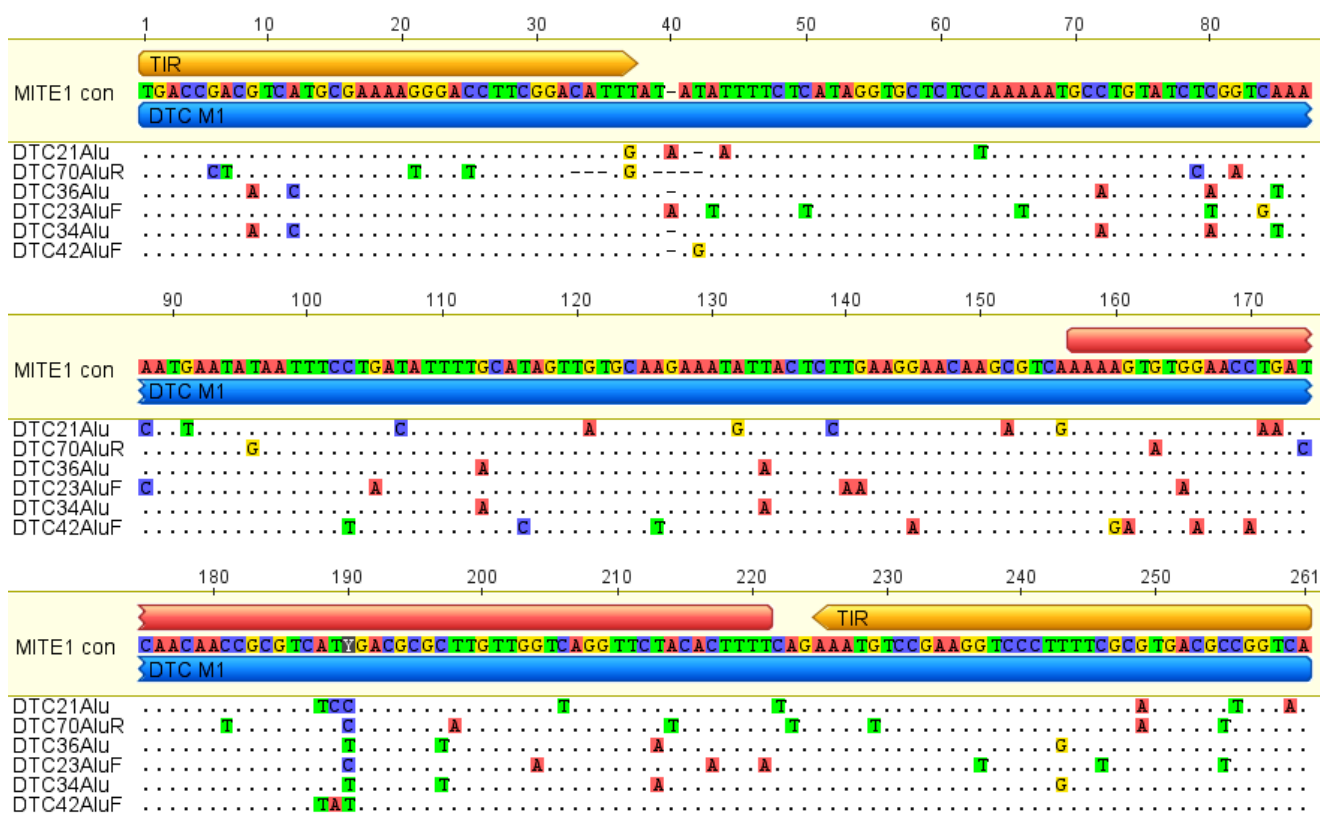
Detaljnijom analizom jednog od ponavljanja uočeno je da se radi o nukleotidnom slijedu dugom 261 pb, na čijim krajevima se nalaze terminalna invertna ponavljanja duljine 37 pb. Ovaj slijed na samim krajevima sadrži dinukleotidni motiv TA. Uzevši u obzir prisutnost TIR slijedova i mogućeg TSD mjesta, zaključeno je da bi ova sekvenca mogla predstavljati MITE element, koji je nazvan DTC M1. Cjeloviti DTC M1 elementi su pronađeni unutar 6 genomskih fragmenata školjkaša *D. trunculus*: DTC21Alu, DTC23AluF, DTC34Alu, DTC36Alu, DTC42AluF, DTC70AluR (Slika 3.20), dok se krnji DTC M1 elementi mogu pronaći kod još 8 fragmenata (DTC47AluF, DTC39Alu, DTC44Alu, DTC1BamF, DTC6BamR, DTC37AluR, DTC30AluR, DTC3BamF), pri čemu su uzeti u obzir samo slijedovi duži od 115 pb.

Unutar cjelovitih elemenata međusobna nukleotidna sličnost varira između 85.1% i 100%, uz prosječnu vrijednost 90,3%. Kod krnjih fragmenata nije uočena nikakva preferencija što se tiče mjesta prekida, a podjednakom učestalošću delecijama je zahvaćen 5' i 3' kraj. Klon DTC42AluF se razlikuje od ostalih po TSD sekvenci, jer je umjesto TA duplikacije na krajevima fragmenta u ovom slučaju prisutan dinukleotid TT. Na poziciji 156-220 elementa DTC M1, neposredno ispred terminalnog invertnog ponavljanja, nalazi se palindromska sekvenca duga čak 65 pb (Slika 3.21). Nije uočena nikakva konzerviranost u nukleotidnim slijedovima koji prethode DTC M1 elementima, kao ni kod onih koji nakon DTC M1 elemenata slijede u genomskom okolišu.

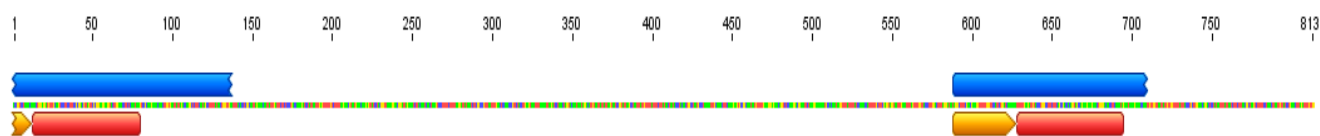
Među klonovima koji sadrže nepotpune DTC M1 elemente zanimljivim se pokazao DTC37AluR. Unutar njega se DTC M1 element nalazi na 2 lokacije, međusobno udaljene ~400 pb i oba puta je prisutan u krnjem obliku (Slika 3.22). Ukoliko se vratimo na Sliku 18b koja prikazuje rezultat *Local blast* pretrage s DTC37AluR klonom, možemo opaziti da prva dva grupiranja sekvenci pripadaju upravo elementu DTC M1 dok treće grupiranje pripada nekom dosad nepoznatom ponavljanju ili mobilnom elementu, koji je svojom ugradnjom mogao dovesti do delecije 3' kraja DTC M1 na kojeg se neposredno nastavlja.



Slika 3.20: 6 genomskih fragmenata s cjelovitim MITE elementima. DTC M1 element je označen plavo, terminalna invertna ponavljanja narančasto, a TA duplikacija na mjestu insercije ljubičasto.



Slika 3.21: Svrnjenje nukleotidnih sljedova DTC M1 elemenata. Terminalna invertna ponavljanja su označena narančasto, a palindromska sekvenca crveno.

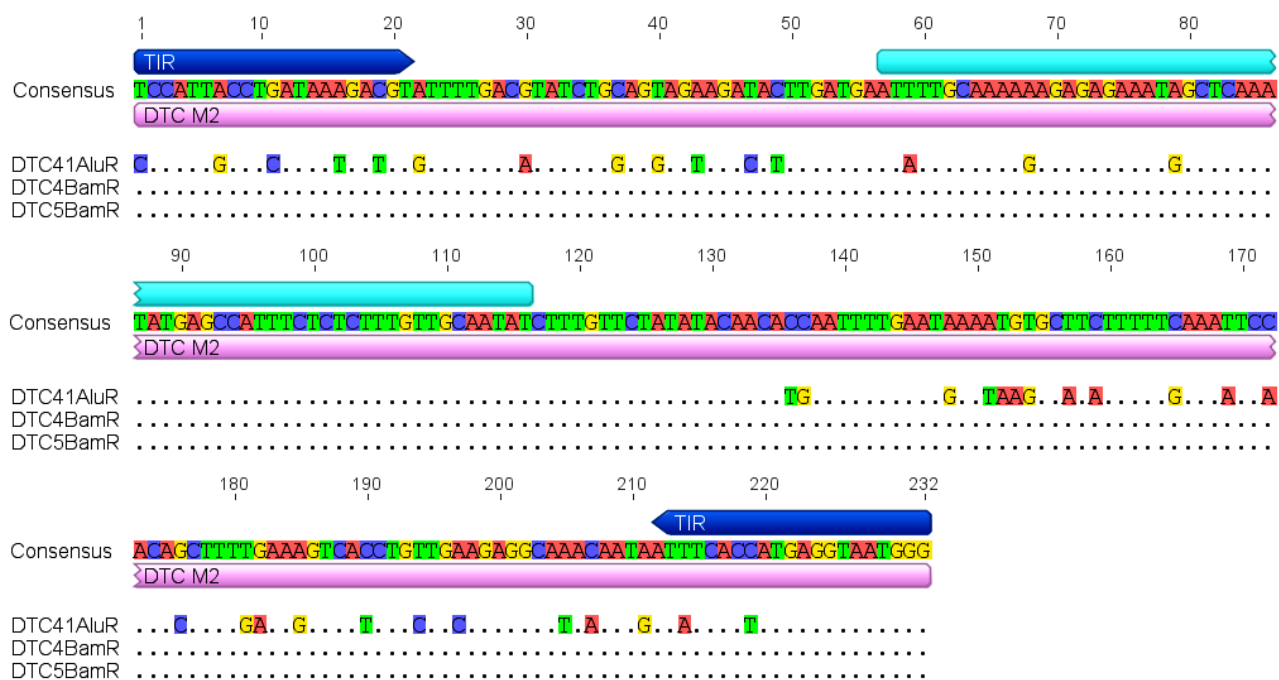


Slika 3.22: Fragmentirani DTC M1 elementi unutar DTC37AluR klona. Sami elementi su označeni plavo, prisutna terminalna invertna ponavljanja narančasto, a palindromske sekvence crveno.

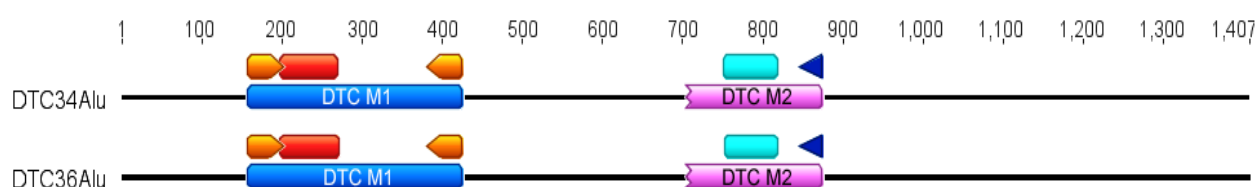
3.1.7. DTC M2

Analizom jednog od sljedećih grupiranja uočeno je da se radi o još jednom MITE elementu, koji je nazvan DTC M2. Radi se o elementu dugom 232 pb, na čijim krajevima se nalaze terminalna invertna ponavljanja duljine 21 pb, dok duplikacije na mjestu insercije nisu uočene. Na poziciji 57-116 elementa DTC M2 nalazi se 60 pb duga palindromska sekvenca. Cjeloviti DTC M2 elementi su pronađeni unutar 3 genomska fragmenta školjkaša *D. trunculus*: DTC41AluR, DTC4BamR, DTC5BamR (Slika 3.23), dok se dijelovi DTC M2 elementa mogu pronaći u još 8 fragmenata (DTC34Alu, DTC36Alu, DTC38Alu, DTC82AluF, DTC82AluR, DTC41AluR, DTC5BamR, DTC7Alu). Krnjim varijantama je zajedničko da redovito imaju prisutnu palindromsku sekvencu, a neke od krnjih varijanti se i sastoje samo od tog nukleotidnog slijeda. Unutar cjelovitih elemenata međusobna nukleotidna sličnost varira između 83,2% i 100%, uz prosječnu vrijednost 88,8%. Nije uočena nikakva sačuvanost nukleotidnih slijedova koji okružuju element DTC M2.

Za klonove DTC34Alu i DTC36Alu je uočeno da sadrže oba MITE elementa, DTC M1 i DTC M2, međusobno udaljena 286 pb, pri čemu je DTC M2 okrnjen za 54 pb (Slika 3.24).



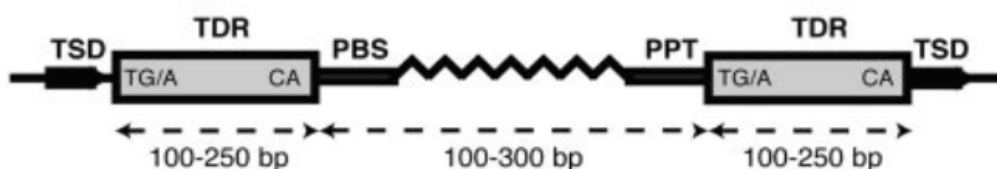
Slika 3.23: Svrnjenje nukleotidnih sljedova DTC M2 elemenata. Terminalna invertna ponavljanja su označena tamno plavo a palindromska sekvenca svjetlo plavo.



Slika 3.24: DTC M1 (plavo) i DTC M2 (ružičasto) elementi unutar DTC34Alu i DTC36Alu klonova.

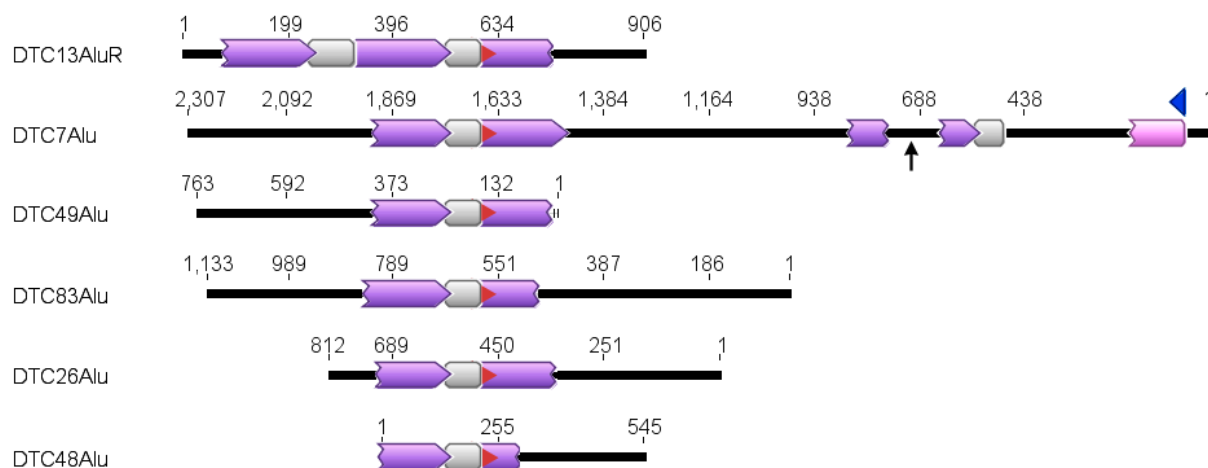
3.1.8. DTC T1

Analizom idućeg grupiranja uočeno je da se radi o elementu TRIM, tipu LTR retrotranspozona, koji je u ovom slučaju nazvan DTC T1. Element je pronađen kod genomskih fragmenata: DTC13AluR, DTC7Alu, DTC49Alu, DTC83Alu, DTC26Alu i DTC48Alu. DTC T1 sadrži TDR motive duljine 202 pb, unutarnju domenu duljine 66 pb, nakon koje slijedi PPT (eng. *Polypurine Tract*) duljine 7 nukleotida. Duplikacije na mjestu insercije nisu uočene, kao ni PBS mjesta. Na Slici 25 je prikazana općenita struktura LTR retrotranspozona porodice TRIM.



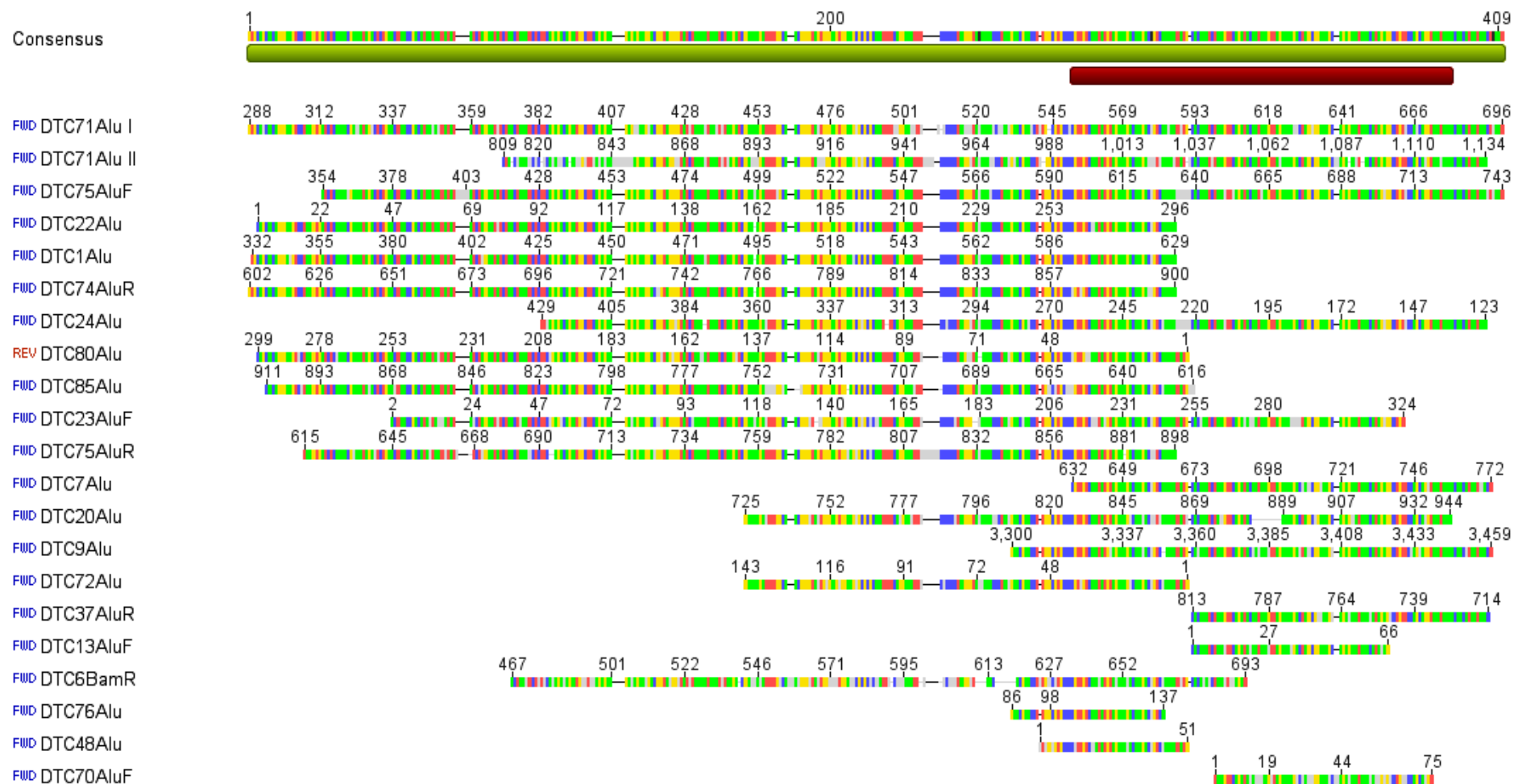
Slika 3.25: Općenita struktura elemenata TRIM. TSD označava duplikacije na mjestu insercije, TDR direktna ponavljanja, PBS vezno mjesto početnice, PPT polipurinski trakt. (Preuzeto iz Witte i sur. 2001)

Unutar ovih 6 klonova TDR domene su uglavnom prisutne u krnjim varijantama (Slika 3.26), a broj deletiranih nukleotida varira od 24 - 53 na 5' kraju te 37 - 115 nukleotida na 3' kraju. Kod unutarnjih domena je također prisutno nekoliko kraćih varijanti, a međusobna nukleotidna sličnost varira između 71,9% i 92,7%, uz prosječnu vrijednost 84,5%. Kraće varijante unutarnjih domena prisutne su između prvog i drugog TDR-a fragmenta DTC13AluR (63 nukleotida); uz treći TDR fragmenta DTC7Alu (53 nukleotida) te kod fragmenta DTC48 (55 nukleotida), kod kojeg je i polipurinski trakt kraći nego kod ostalih. Nije uočena nikakva konzerviranost u nukleotidnim slijedovima koji prethode DTC T1 elementima, kao ni kod onih koji u genomskim sekvencama slijede nakon njih.



Slika 3.26: Genomski fragmenti koji sadrže elemente DTC T1. TDR-ovi su označeni ljubičasto, unutarnje domene sivo, polipurinski trakt crvenim strelicama. Element DTC M2 je označen ružičasto, njegovo terminalno invertno ponavljanje plavo. Sekvenca insertirana unutar dislocirane TDR domene fragmenta DTC7Alu je označena strelicom.

Osim uobičajene strukture TRIM elemenata, pojava onih koji sadrže samo jednu ili onih s nekoliko TDR domena nije neuobičajena. Kod fragmenta DTC7Alu jednostruka TDR domena je na poziciji 629 – 769 prekinuta s insercijom nukleotidne sekvence duge 140 pb (Slika 3.26), uz pojavu TGA duplikacije na mjestu insercije. *Local blast* pretraga pokazuje da je slijed duljine 140 pb sastavni dio dosad nepoznate raspršeno ponovljene sekvence duljine 409 pb koja se pak pojavljuje unutar 20 klonova školjkaša *D. trunculus* (Slika 3.27). *NCBI blast* i pretraga baze podataka *Repbase* međutim ne nalaze nikakvu značajnu sličnost s tom sekvencom. U sklopu fragmenta DTC7Alu uočena je i krnja varijanta MITE elementa DTC M2. Ovoj varijanti nedostaje 117 pb, tako da element počinje direktno s palindromskom sekvencom.



Slika 3.27: Dosad nepoznata repetitivna sekvenca prisutna u velikom broju klonova školjkaša *D. trunculus*. tamno crveno je označeno područje sekvence koje se nalazi insertirano unutar treće TDR domene genomske fragmenta DTC7Alu.

3.2. Pregled repetitivnih sekvenci u genomu školjkaša *Ruditapes decussatus*

U svrhu detekcije i karakterizacije repetitivnih sekvenci genoma školjkaša *Ruditapes decussatus*, također je korištena metoda analize otisaka kolonija u hibridizaciji s cjelokupnom genomskom DNA, parcijalno pocijepanom s restriksijskom endonukleozom *AluI*. Da bi se utvrdilo postoji li sličnost i preklapanje u repetitivnim dijelovima genoma kod srodnih vrsta *R. decussatus* i *R. philippinarum* ponovljen je isti postupak, ali uz korištenje genomske sonde jedne vrste na genomskoj biblioteci druge vrste. Nakon izolacije i utvrđivanja nukleotidnog slijeda, dobiveni genetički fragmenti od interesa uspoređeni su s bazom podataka *NCBI GenBank* te putem pretraživača *CENSOR* s bazom *Repbase* koja sadrži kolekciju različitih tipova poznatih repetitivnih sekvenci i mobilnih elemenata. Dobiveni rezultati prikazani su u Tablici 3.3. Klonovi školjkaša *R. decussatus* identificirani s genomskom sondom vlastite vrste nose oznaku „D“, a klonovi školjkaša *R. decussatus* identificirani genomskom sondom vrste *R. philippinarum* nose oznaku „DP“. Iako su u pretraživanju biblioteke fragmenata školjkaša *R. decussatus* s hibridizacijskom sondom vrste *R. philippinarum* 4 klona dala pozitivan signal, nije uočen niti jedan koji bi hibridizirao s obje sonde.

Veličina haploidnog genoma školjkaša *R. decussatus* iznosi $1,77 \times 10^9$ pb prema *Animal Genome Size Database*, pri čemu izolirani klonovi obuhvaćaju oko 28 200 pb. Klonirani inserti su duljine 41 pb – 4500 pb, a prosječan GC sastav fragmenata iznosi 35.0%. Slično kao i u analizi genomske DNA školjkaša *D. trunculus*, dobiveni sljedovi mogu se kategorizirati u 7 skupina s obzirom na različite tipove repetitivne DNA koje sadrže. Prva skupina obuhvaća genomske fragmente koji sadrže isključivo uzastopno ponovljene monomere satelitne DNA, pri čemu je kod klonova D17 i D19 nađena satelitna DNA BIV160, dok je kod D32 nađena satelitna DNA DTHS3. Druga skupina sadrži monomere satelitne DNA DTHS3 vezane uz nerepetitivne sljedove genoma školjkaša (D26, D7F, DP36). Sljedeće 2 grupe fragmenata su pokazale sličnost s različitim tipovima mobilnih elemenata, jedna s različitim porodicama DNA transpozona (klonovi D1, D3, D8, D13, D18, D20, D27), a druga s porodicom non-LTR retrotranspozona (klon D9). Peta skupina sekvenci obuhvaća one koje sadrže više od jednog tipa repetitivne DNA, a u njihovom se sastavu mogu pronaći različite kombinacije elemenata prisutnih u svim prethodnim skupinama. Ovdje se nalaze klonovi: D2R, D10, D12, D14, D17, D19, DP34. Šesta skupina sadrži sljedove koji pokazuju sličnost s retroviralnim sekvencama (D7R, D27).

klon	Satelitna DNA		Transpozon			cDNA	ostalo	Duljina klona
	Samo satelitna DNA	Satelitna DNA prisutna uz neku drugu sekvencu	DNA transpozon	Retrotranspozon				
				LTR	Non-LTR			
D1			209-294 Helitron-N9_CGi			128-315		324
D2F						665-876		919
D2R			34-190 Polinton-1_HM	225-262 DIASPORA_I		396-498		912
D3			1222-1406 Helitron-N9_CGi			1132-1393	1270-1400 S. Solidiss. gen za cink finger protein	1411
D4								168
D5								199
D6F						69-146, 833-919		919
D6R								913
D7F		571-646 DTHS3				236-406, 567-654		918
D7R						475-738	149-220 ETnERV3 Endogeni Retrovirus	911
D8			106-146 HELITRONY2			122-207		533
D9					16 -143 ISABELLE_TM		11-223 mRNA za teški lanac miozina	230
D10			170-222 CvA	274-430 Gypsy157-I_DR 640-672 Copia-40_ST-I		89-236, 729-802	228-605 Pol poliprotein transpozona 17.6; Reverzna transkriptaza	803
D11						1-56		61
D12		149-221, 485-552			410-482	227-406, 554-806		1003

		DTHS3			SINE-4_CQ			
D13			84-161 Helitron-N36_CGi			1-74		184
D14			1298-1394 Mariner-19_HM	670-714 Copia-65_SB-I		1635-2124	1567-1911 helitron 3, helitronu-sličan transpozon replikaza/helikaza/endonukleaza	2124
D15						269-344, 422-919		243
D16						659-776, 660-802	66-299, 368-457 LINE-1 retrotranspozonu sličan element	1249
D17	BIV160		6-796 Helitron-N2d_CGi			1-837		837
D18			824-898 P-1_AP					925
D19	BIV160		43-655 Helitron-N2d_CGi			1-660		660
D20			916-1116 Helitron-N2d_CGi			624-1685	2x290 pb dugi nizovi minisatelitnih ponavljanja 667-782, 988-1250 S. solidissima cnRNA154	1720
D21								292
D22						204-288		294
D23								83
D24								52
D25						240-514	233-514 S. solidiss. gen za cink finger protein	958
D26		351-545 DTHS3				348-544	363-541 S. solidiss. gen za cink finger prot.	545
D27			552-636 Helitron-N9_CGi			313-657	246-293 ZFERV-2-I_DR Endogeni Retrovirus	1026
D28						1-288		289
D29								90
D30								502

								Rezultati
D32	DTHS3					1-142	2-142 S. solidiss. gen za cink finger prot.	145
DP33						85-403	117-217 Msx gen za Msx protein	432
DP34			12-104 EnSpm11_SB		446-585 PNL_SM		1156-1431 dio 18S rRNA gena	1431
DP35								41
DP36		1-149 DTHS3				2-149	34-148 S. solidiss. gen za cink finger prot.	224

Tablica 3.3: Pregled dobivenih genetičkih fragmenata s obzirom na prisutnost različitih repetitivnih elemenata. Brojevi označavaju nukleotidnu poziciju satelitne DNA ili mobilnog elementa unutar klona, cDNA kolona daje pregled odsječaka kojima su pronađeni transkribirani ekvivalenti.

Rezultati prikazani u cDNA koloni Tablice 3.3, dobiveni pretragom EST kolekcije unutar baze *NCBI GenBank*, pokazuju da se unutar velikog broja klonova dobivenih ovom metodom nalaze područja koja su vjerojatno transkripcijski aktivna. Ta se područja kod 11 klonova poklapaju s onim istim djelovima kod kojih je uočena sličnost s nekim mobilnim elementom ili satelitnom DNA, dok je kod nekih uočena sličnost i s konkretnim proteinskim produktom (vidjeti kolonu „ostalo“ Tablice 3.3).

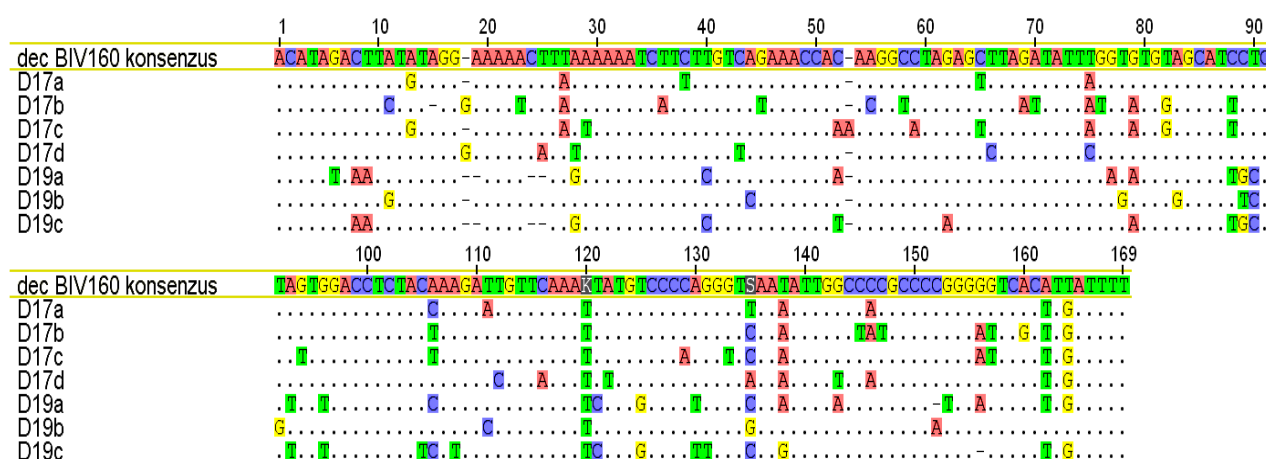
Tablica 3.4 sumarno prikazuje osnovne tipove mobilnih elemenata (uz izdvojene glavne porodice) s kojima je uočena sličnost unutar genomskih fragmenata školjkaša *R. decussatus*.

Tip elementa
DNA transpozon (13)
EnSpm
Helitron
Mariner/Tc1
P
Polinton
Retrovirus (2)
ERV1
ERV2
LTR Retrotranspozon (4)
Copia
Gypsy
Non-LTR Retrotranspozon (3)
L1
Penelope
SINE
Mobilni elementi (22)

Tablica 3.4: Osnovni tipovi mobilnih elemenata koji su pokazali sličnost s genomskim fragmentima školjkaša *R. decussatus*, uz izdvojene glavne porodice elemenata.

3.2.1. Satelitne DNA izolirane iz vrste *Ruditapes decussatus*

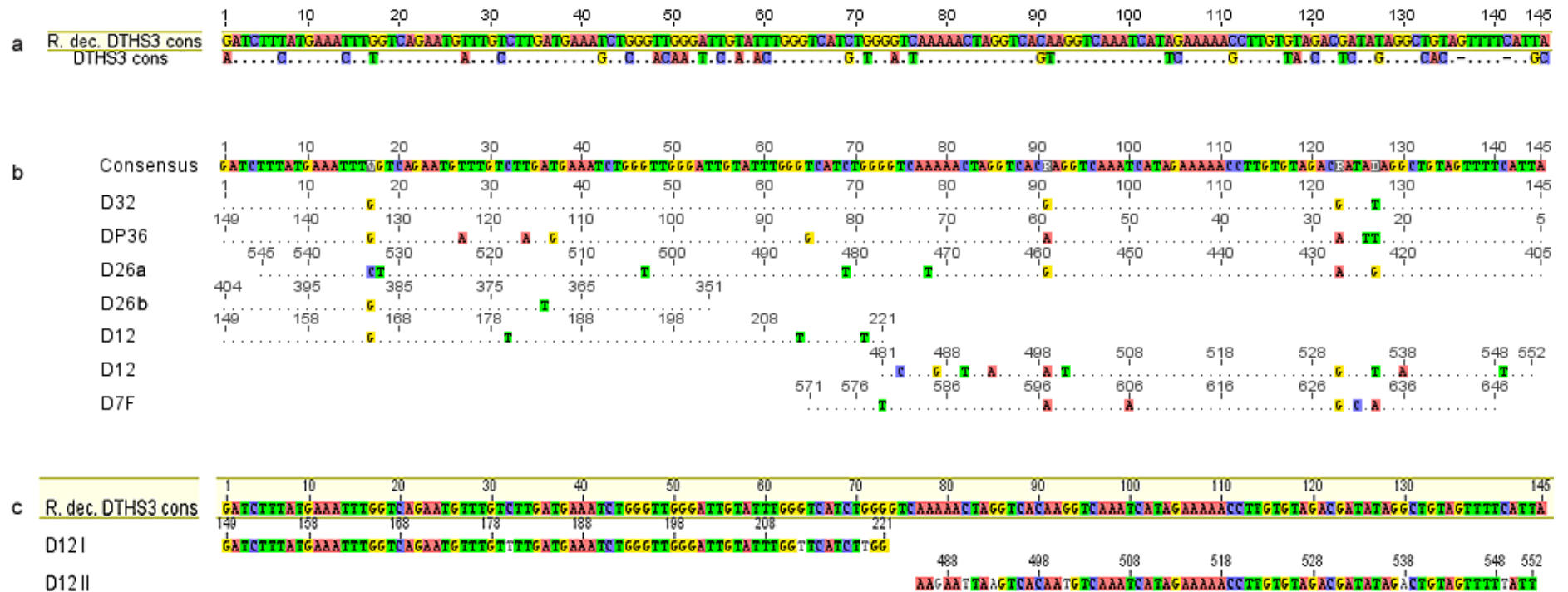
Unutar prve skupine, koja obuhvaća genomske fragmente sastavljene isključivo od uzastopno ponovljenih monomera satelitne DNA BIV160, nalaze se klonovi D17 i D19, dugi 837 i 660 pb. Nakon izdvajanja monomernih jedinica i njihove analize (Slika 3.28), utvrđeno je da im međusobna nukleotidna sličnost varira u rasponu od 75,7% - 92,7%, sa srednjom vrijednosti 82,8%. Dobiveni rezultati su u skladu s ranije opaženim vrijednostima za monomere satelitne DNA BIV160 izolirane iz vrste *R. decussatus* metodom lančane reakcije polimerazom (Plohl i sur. 2010), pri čemu je bila uočena nukleotidna sličnost od 75,6% - 96,4%, sa srednjom vrijednosti 86,4%.



Slika 3.28: Svrnjenje monomera izdvojenih iz D17 i D19 klonova uz usporedbu s konsenzus sekvencom satelitne DNA BIV160 iz vrste *R. decussatus*.

Posljednji klon iz prve skupine, D32, zajedno s 4 klona iz druge skupine, sadrži monomere DTSH3 satelitne DNA koji se nastavljaju na neke druge genomske sljedove ovog školjkaša. Iako konsenzus sekvenca DTSH3 monomera pronađenih u vrsti *R. decussatus* dijeli 73,1% homologije s konsenzus sekvencom DTSH3 monomera izvorno pronađenih kod školjkaša *D. trunculus* (Slika 3.29a), ipak se radi o vrsno-specifičnoj skupini unutar koje je visoka homologija. Ona u prosjeku iznosi 93,9%, uzevši u obzir cjelovite monomerne jedinice iz svrnjenja vidljivog na Slici 29b. Monomer satelitne DNA DTSH3 prisutan kod klona D12 je prekinut na sredini (Slika 3.29c) i to insercijom 263 pb duge nukleotidne sekvence. *NCBI*

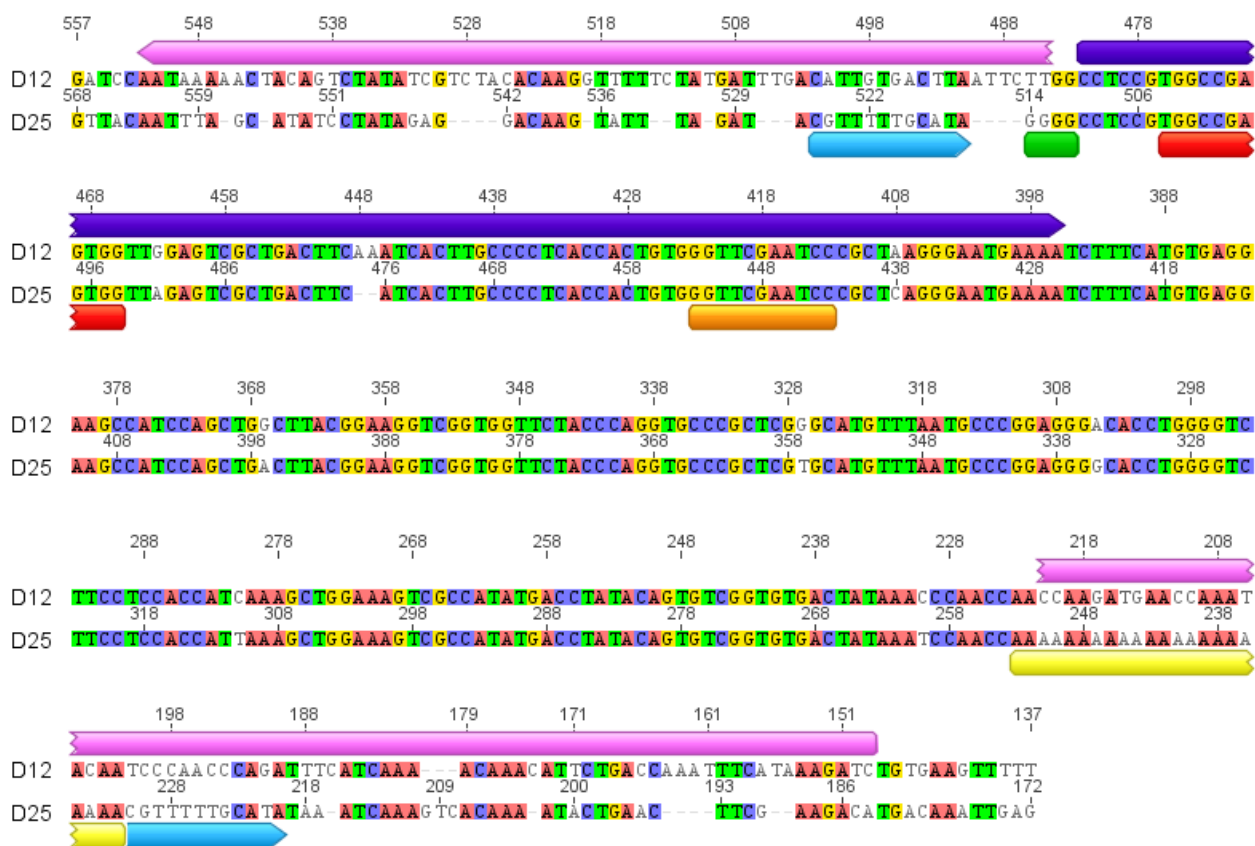
blast i pretraga baze podataka *Repbase* ne nalaze nikakvu značajnu sličnost s insertiranom sekvencom, no *Local blast* pokazuje da se radi o nukleotidnom slijedu koji se također nalazi i u sklopu klona D25.



Slika 3.29: Svrnjenje konsenzus sekvenci satelitne DNA DTSH3 školjkaša *R. decussatus* i od ranije poznate satelitne DNA DTSH3 školjkaša *D. trunculus* (a). Svrnjenje izdvojenih monomera satelitne DNA DTSH3 pronađene u vrsti *R. decussatus* (b). Prekinut monomer satelitne DNA DTSH3, nađen u klonu D12 (c).

3.2.2. Element RUDI 1

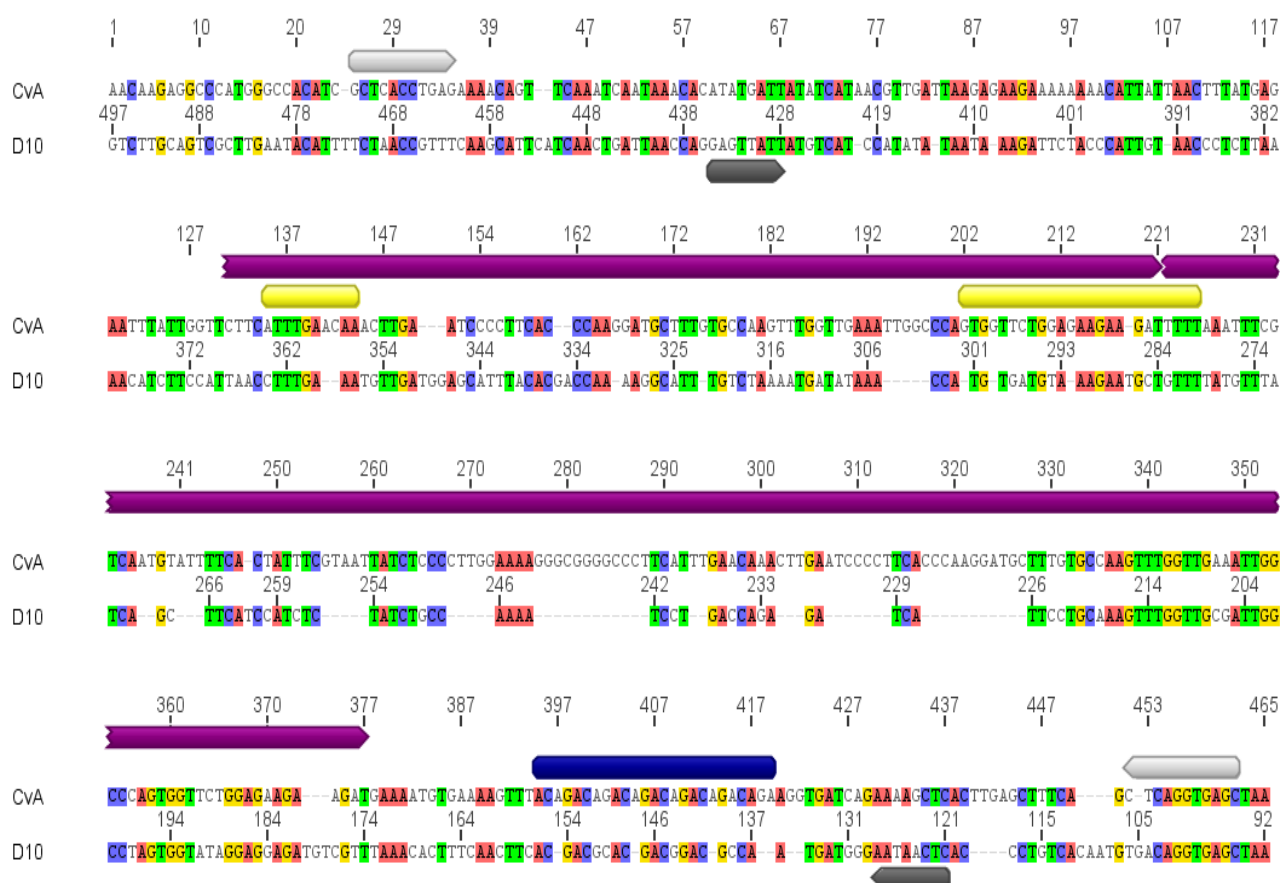
Detaljnijoj analizom sekvence inserirane u monomer satelitne DNA DTHS3 utvrđeno je da se radi o retrosekvenci tipa SINE. Usporedbom elemenata izdvojenih iz klonova D12 i D25, opaža se visok stupanj homologije među njima, 96,6% na području od 263 pb. Pretraga u sklopu *tRNAscan-SE* servisa je pokazala da je 5' kraj elementa porijeklom od stanične tRNA, a sačuvani su i box A i box B motivi promotora RNA polimeraze III (Slika 3.30). Kod klona D25 uočljive su i 12-nukleotidne duplikacije na mjestu insercije, kratki G-bogati trakt na 5' kraju elementa kao i 22 nukleotida dug poli A rep. Ove 3 navedene strukturne karakteristike međutim izostaju kod klona D12, kod kojeg je SINE element RUDI inseriran usred monomera satelitne DNA.



Slika 3.30: Svrnjenje nukleotidnih sljedova elementa RUDI nađenih u sklopu klonova D12 i D25. Duplikacije na mjestu insercije označene su svijetlo plavo, kratki G-bogati trakt zeleno, 5' kraj elementa koji je tRNA porijekla ljubičasto, box A crveno, box B narančasto, poli A rep žuto, a dijelovi monomera DTHS3 satelitne DNA ružičasto.

3.2.3. Element D10

Unutar klona D10 uočeno je postojanje nukleotidnog sljedaja duljine 406 pb, za koji je opaženo da pokazuje strukturnu, a djelomično i nukleotidnu sličnost s prethodno spomenutim mobilnim elementom CvA (Gaffney i sur. (2003). Osim u jednoj regiji više homologije, podudarnost u nukleotidnoj sekvenci je cijelom dužinom podjednaka i relativno niska (53,8%). Osnovna strukturna građa elementa je dobro očuvana (Slika 3.31). Međutim, dok se invertna ponavljanja kod elementa CvA nalaze periferno, kod elementa D10 su pomaknuta 20-ak parova baza prema centralnom dijelu. Mikrosatelitna regija elementa CvA sastoji se od ACAG ponavljanja dok se kod elementa D10 nalazi kombinacija tetranukleotidnih ACGA i trinukleotidnih CGC ponavljanja. Konzervirane regije prisutne kod CvA elementa i cijelog niza repetitivnih sekvenci (Slika 3.4) u ovom slučaju nisu očuvane, a opažen je i niz delecija unutar jedinica centralnih ponavljanja.



Slika 3.31: Svrnjenje nukleotidnih sljedova dijela elementa CvA i elementa D10. Svijetlo sivo i tamno sivo označava invertna ponavljanja, ljubičasto centralna ponavljanja, tamno plavo mikrosatelitne sekvence, a žuto je označeno područje konzerviranih blokova koje element CvA dijeli s nekoliko porodica satelitnih DNA.

3.3. Pregled repetitivnih sekvenci u genomu školjkaša *Ruditapes philippinarum*

U svrhu detekcije i karakterizacije repetitivnih sekvenci genoma školjkaša *Ruditapes philippinarum* korištena je ista metoda kao i kod prethodno opisanih vrsta. Da bi se utvrdilo postoji li sličnost i preklapanje u repetitivnim dijelovima genoma školjkaša *R. decussatus* i blisko srodne vrste *R. philippinarum*, primjenjen je isti eksperimentalni pristup kao i ranije, hibridizacija uz korištenje genomske sonde jedne vrste na genomskoj biblioteci druge vrste. Nakon izolacije i utvrđivanja nukleotidnog slijeda, dobiveni genetički fragmenti od interesa ponovno su uspoređeni s bazom *NCBI GenBank* te putem pretraživača *CENSOR* s bazom poznatih repetitivnih sekvenci i mobilnih elemenata *Repbase*. Dobiveni rezultati prikazani su u Tablici 3.5. X klonova dobivenih genomskom sondom *R. philippinarum* nose oznaku „P“. Oni dobiveni genomskom sondom *R. decussatus* na genomskoj biblioteci vrste *R. philippinarum* označeni su „PD“ (19 klonova). S obje sonde hibridiziralo je čak 16 klonova (P14, P16, P18, PD22, P31, P43, P44, P45, P52, PD53, PD55, PD56, PD58, PD59, PD60, PD61).

Veličina haploidnog genoma ovog školjkaša iznosi $1,93 \times 10^9$ pb prema *Animal Genome Size Database*, pri čemu izolirani klonovi obuhvaćaju oko 51 100 pb. Duljina kloniranih fragmenata je u rasponu od 32 pb – 6000 pb, a prosječan GC sastav fragmenata iznosi 36,5%. Dobiveni sljedovi se mogu kategorizirati u 6 skupina s obzirom na različite tipove repetitivne DNA koje sadrže. Prva skupina obuhvaća genomske fragmente koji sadrže isključivo uzastopno ponovljene monomere satelitne DNA *phBgIII400* ranije opisane u vrsti *R. philippinarum* (Passamonti i sur. 1998), prisutne unutar klonova P4, P5, P17, P47, P49. Druga skupina sadrži monomere satelitne DNA *DTHS3* prethodno opisane u školjkašu *D. trunculus* (Plohl i Cornudella 1997), u ovom slučaju vezane uz nerepetitivne i zasad neokarakterizirane sljedove genoma, pri čemu je kod klonova P12, P13, P15, P34, P35, P41, P42, P43, PD60 u pitanju satelitna DNA *phBgIII400*, a kod fragmenata PD53F i PD53R radi se o satelitnoj DNA *DTHS3*. U klonu P46 prisutne su sekvence obje satelitne DNA. Klon P18 sadrži pak kombinaciju satelitne DNA *phBgIII400* i satelitne DNA *BIV160* (Plohl i sur. 2010). Sljedeće 3 grupe fragmenata su pokazale sličnost s različitim tipovima mobilnih elemenata. Sličnost s različitim porodicama DNA transpozona uključuje genomske fragmente P7, P9, P11, P14F, P19, PD28, P45R, P48, P52 i PD56F sličnost s LTR retrotranspozonom uključuje genomske fragmente P6R, PD 22, PD24, PD26, P44R, P45F i P50F, dok sličnost s non-LTR retrotranspozonom imaju P37 i P50R.

Šesta skupina sekvenci obuhvaća one koje sadrže više od jednog tipa repetitivne DNA, a u njihovom se sastavu mogu pronaći različite kombinacije elemenata prisutnih u svim prethodnim skupinama. Ovakve sekvence sadrže genomske fragmenti P1, P2, P3, P16, P18, PD23, PD27F, P44F, P46 i PD53F.

Rezultati u sklopu cDNA kolone Tablice 3.5, dobiveni pretragom kolekcije EST (eng. *expressed sequence tag*) unutar baze podataka *NCBI GenBank Database*, u velikom broju klonova pokazuju područja koja su vrlo vjerojatno transkripcijski aktivna. Ta područja se kod 9 genomskih fragmenata poklapaju s onim istim dijelovima kod kojih je uočena sličnost s nekim mobilnim elementom ili satelitnom DNA.

Tablica 3.6 sumarno prikazuje osnovne tipove mobilnih elemenata (uz izdvojene njihove glavne porodice) s kojima je uočena sličnost unutar *R. philippinarum* genomskih fragmenata.

klon	Satelitna DNA		Transpozon		cDNA	ostalo	Duljina klona	
	Samo satelitna DNA	Satelitna DNA prisutna uz neku drugu sekvencu	DNA transpozon	Retrotranspozon				
				LTR				Non-LTR
P1		338-658 <i>phBglII400</i>	168 -282 Helitron-N2_SP		134-285		663	
P2		311-718 <i>phBglII400</i> 13-55 DTRS sat		525-602 Copia-104_GM-I	10-172		718	
P3		1-223 <i>phBglII400</i>	500 -691 Ginger1-11_HM			775-1162 12 puta uzastopce ponovljena 19-nukleotidna ponavljanja	1642	
P4	<i>phBglII400</i>						403	
P5	<i>phBglII400</i>						97	
P6F					255-916		916	
P6R				465-507 SZ-37_I	67-272, 761-824, 819-911		911	
P7				331-653 Gypsy-72_CGi-I 727-840 Gypsy-1_ARO-I	1-175, 1015-1159	1156-1217 dio 28S rRNA gena	1217	
P8							134	
P9			10-100 hATx-22_SM				107	
P10					127-358		358	
P11				331-653 Gypsy-72_CGi-I 728-841 Gypsy-1_ARO-I	1-175, 854-1161	1158-1219 dio 28S rRNA gena 162-563 retr-1 retrotranspozonu sličan element	1219	
P12		1-316 <i>phBglII400</i>					656	
P13		1-316 <i>phBglII400</i>					656	

P14F			637-732 DNA8-96_AP 773 -848 Helitron-N3_SP			158-416		916
P14R								911
P15		1-316 phBglII400						656
P16		143-241 phBglII400	1-83 Helitron-N1_SP			6-148		241
P17	phBglII400							97
P18		770-1083 phBglII400 , 1890-1985 phBglII400 , 2008-3433 BIV sat	3610-3688 EnSpm-5_VV		130-204 L1-56_XT	2132-3432	1116-1583 reverzna transkriptaza iz mobilnog elementa <i>jockey</i> 1083-1577 Pol protein	3995
P19			134-224 Polinton-2A_NV			223-277		509
P20								136
P21						127-358		358
PD22				331-653 Gypsy-72_CGi-I 727-840 Gypsy-1_ARO-I		1-175, 854-1161	1156-1217 dio 28S rRNA gena	1217
PD23			193 -253 DNA8-71B_AP		332-366 CR1-4_DK	116-246		568
PD24				125 -177 ATHILA-7_SBi-LTR				222
PD25								65
PD26				69 -211 Gypsy9-SM_LTR				292
PD27F			583-630 OposCharlie3b		214-426 SIRE3_TC			923

PD27R								917
PD28			274-399 Ginger1-5_HM				376-409 poli A	457
PD29								242
PD30								80
P31							Mikrosatelitna ponavljanja GCA x 9, AGT x 6, TGT x 3, AGT x 14	291
P32								75
P33								168
P34		32-128 <i>phBglII400</i>						128
P35		2-105 <i>phBglII400</i>						145
P36						20-118		118
P37					1-83 CR1-2_BF	81-200	2-88 endonukleaza/rev. transkriptaza	200
P38						1-45		155
P39						1-168		220
P40								83
P41		32-128 <i>phBglII400</i>						128
P42		153-251 <i>phBglII400</i>				45-159		251
P43		332-428 <i>phBglII400</i>				1-136		428
P44F			617-699 DNA-TA-12_CQ	492-547 GYPHY17-I_AG		1-287, 409-569	800-810 poli A	922
P44R				92-201 Gypsy4-I_Dmoj				911
P45F				513-600 BEL-18_DWil-I 737-786 Gypsy-10_OD-I		1-169		920
P45R			134-167 Rehavkus-1_TC 453-576 DNA-11N_SBi			757-831		913

P46		1-177 DTHS3 1906-2004 phBglII400	147-256 CVA 1008-1168 Helitron-1_DPe		1217-1266 Crack-1_NV	1-177, 1757-1916	15-177 Spisula solidissima gen za cink finger protein Sso-Zic	2004
P47	phBglII400							143
P48			14-109 Helitron-N2_CGi			1-67		115
P49	phBglII400							307
P50F				373-415 SZ-37_I		1-282, 608-855		916
P50R					544-718 CR1-2_NV 721-888 CR1-41_BF		514-891 RNA-dirigirana DNA polimeraza iz mobilnog elementa jockey	912
P51								62
P52			303-370 HELITRONY1C			12-313	174-317 Msx gen	1121
PD53F		1187-1660 DTHS3	1582-1628 Helitron-N9_CGi			609-659, 951-1012, 1189- 1666	1187-1661 Spisula solidissima gen za cink finger protein Sso- Zic	1666
PD53R		1-687 DTHS3				1-685		911
PD54						1-32		32
PD55								188
PD56F			43-148 Transib-18_HM 284-342 HELMET2					919
PD56R						166-394		913
PD57								117
PD58								146
PD59F						707-834		921
PD59R							733-769 poli A	915
PD60		1-106 phBglII400				102-149		160
PD61								58
PD62						194-434		469

Tablica 3.5: Pregled dobivenih genetičkih fragmenata s obzirom na prisutnost različitih repetitivnih elemenata. Brojevi označavaju nukleotidnu poziciju satelitne DNA ili mobilnog elementa unutar klona, cDNA kolona daje pregled odsječaka kojima su pronađeni transkribirani ekvivalenti.

Tip elementa
DNA transpozon (26)
Academ
EnSpm
Ginger1
Helitron
MuDR
Polinton
Sola
Transib
hAT
LTR Retrotranspozon (11)
BEL
Copia
DIRS
Gypsy
Non-LTR Retrotranspozon (7)
CR1
Crack
Tx1
Mobilni elementi (44)

Tablica 3.6: Osnovni tipovi mobilnih elemenata koji su pokazali sličnost s genomskim fragmentima školjkaša *R. philippinarum*, uz izdvojene glavne porodice elemenata.

3.3.1. Satelitne DNA izolirane iz vrste *Ruditapes philippinarum*

Iz genomskih fragmenata koji su svrstani u prvu i drugu skupinu izdvojeno je 20 nukleotidnih sljedova koji pripadaju satelitnoj DNA *phBgIII400*. Duljina izdvojenih fragmenata satelitnih monomera varira od 97 - 408 pb, a nukleotidna sličnost u odnosu na ranije poznatu konsenzus sekvencu je u rasponu od 82,8% - 96,7%. AT sastav ne odudara od ranije uočenih 60%. Uzevši u obzir da je većina izdvojenih fragmenata značajno kraća od duljine monomera koji iznosi 405 pb, nije se ulazilo u detaljnije analize.

Sljedeća porodica satelitnih DNA, BIV160, pronađena je samo unutar klona P18 i čini 1425 pb dugački niz iz kojeg su izdvojene 2 krnje i 7 cjelovitih jedinica ponavljanja. Njihova duljina, u rasponu od 163 - 169 pb, svrstava ih u „dulju“ varijantu satelitnih monomera ranije pronađenih u vrsti *R. philippinarum* (duljina monomera prve varijante iznosi 157 pb, a druge ~ 166 pb; Plohl i sur. 2010). Unutar dulje varijante, već je ranije bio uočen široki raspon sličnosti jedinica ponavljanja, s najnižom vrijednosti od 43%, dok je sličnost monomera izdvojenih ovom metodom u rasponu 61,5% - 88,6%, uz srednju vrijednost od 70,3%. Ranije uočeni AT sastav od 65,6% je nepromijenjen.

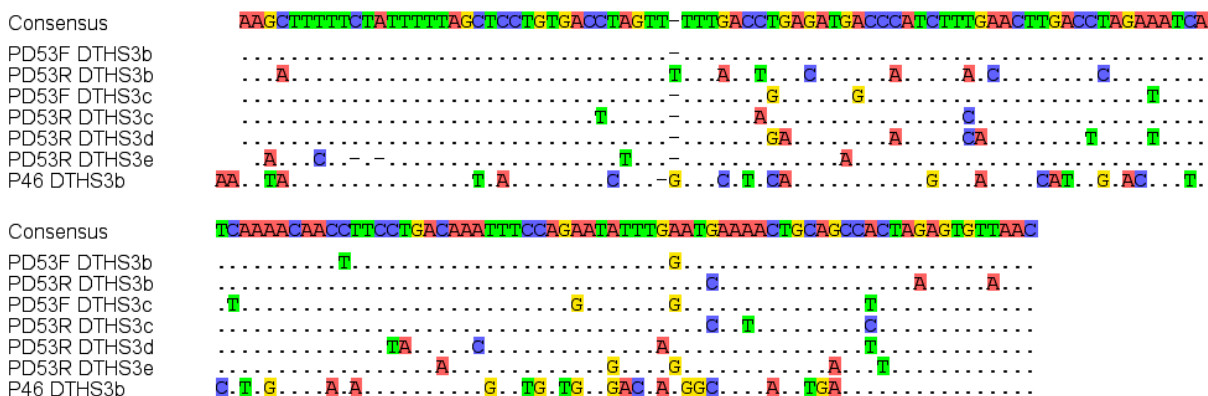
cons	MCAT-----ATAERTATA TAGTAAAAAC-----T-TAAAAAAATCTTCCTGTCAGAAATYACAAGGCCTTAGAGCTTAGATATTGG-ATATGTAACAAT
P18 - BIVi	...--T-----AC.C-----G-----C-----G..TC.T...TTGGAG.AAG-...C-----T.AC...A..TG
P18 - BIVa	...--T-----T.A-----A.T-----T-----G..AG.CCC.A...-A-----T.-CG.AAGG...
P18 - BIVg	..C.-T-----G..A..-----G-----C.TT...TTG..A.G...G..A...A..C..AG..T
P18 - BIVc	..C.-T-----G..-----CT-----AA...AAA...C.AGC...TTG..A.A...G...G..T.-C..A....
P18 - BIVe	..C.-T-----G..-----A-----A...C.TAG.TTG..A.G...G..G...T.-C..A....
P18 - BIVb	..CTAT-----G.AC-----T-----A-----T.C-----T..GGT...-T..G..T...-GTT
P18 - BIVf	..CTATGTTT...CA.G-----T-----A-----T..GGT...-T...T...A...T
P18 - BIVd	..A.-T-----G.AC-----T-----T..C.A-----TT..GGT...-T...TA...T-G.A.G.T
P18 - BIVh	..T-----G.AC-----AAAAAAA------T-----T..GGT...-A...T..G..C...A..T
cons	CWCTARGGANCCTACCAAGWTTGTTCAAAATTAATGCTCTAGTGTAAAAATTTGGCCCCGCCCGGGGGT---CATARKTTTTT
P18 - BIVi	TG.T.GG...C..GCTT...A....
P18 - BIVa	TG...GG.A..C...AA...A..C..T.G...C...TGA...
P18 - BIVg	TGCGA.G.AA.C...AA...A..C..G...T...A...GGT...TGA...
P18 - BIVc	TG.G.G.A.T...AA...A..C..G.G...T.T...TGA...
P18 - BIVe	TG.G.G.A.C...G..A...A...A..C..G.G...A..C...TGA...
P18 - BIVb	TA...TA..TA..A...T.GC..A..T...TA...GG
P18 - BIVf	A...TG..TC...T.AG...TT...GC.C...T.ATT...T...A.A...-G..GT...
P18 - BIVd	A...TA..TA...AGC...T.AC...T...G.A...A..A...GGTCAAT.CTGG...
P18 - BIVh	A...TA..TC...AG...T.GC.C...-T...C..T...T.G.AA.A...-TGA..C

Slika 3.32: Svrnjenje monomera izdvojenih iz klona P18 uz usporedbu s konsenzus sekvencom dulje varijante BIV160 satelitne DNA vrste *R. philippinarum*. Konsenzus sekvencu prikazana je punim nukleotidnim slijedom, dok su kod pojedinih jedinica ponavljanja prikazane samo razlike u odnosu na konsenzus.

Satelitna DNA DTHS3 prisutna je unutar klona P46 te unutar fragmenata PD53F i PD53R. Ova dva fragmenta pripadaju istom klonu PD53, ali nukleotidni slijed cjelovitog inserta nije bio utvrđen. Fragment PD53F je duljine 1666 pb, od čega posljednjih 473 pb sadrži satelitnu DNA DTHS3, baš kao i početnih 687 pb u fragmentu PD53R, duljine 911 pb (Slika 3.33). Za pretpostaviti je da i nesekvencirani centralni dio inserta, a čija ukupna duljina iznosi ~3 kb, također sadrži monomere ove satelitne DNA. Nakon izdvajanja cjelovitih jedinica ponavljanja (Slika 3.34), utvrđeno je da njihova nukleotidna sličnost varira između 67,6% - 95,2% sa srednjom vrijednosti od 83,6%. Ukoliko se uzmu u obzir samo monomeri iz klona PD53 koji, predvidivo, pripadaju istom nizu, tada srednja vrijednost iznosi 89,6%.



Slika 3.33: Shematski prikaz PD53F i PD53R fragmenata klona PD53. Označene su jedinice ponavljanja satelitne DNA DTHS3.

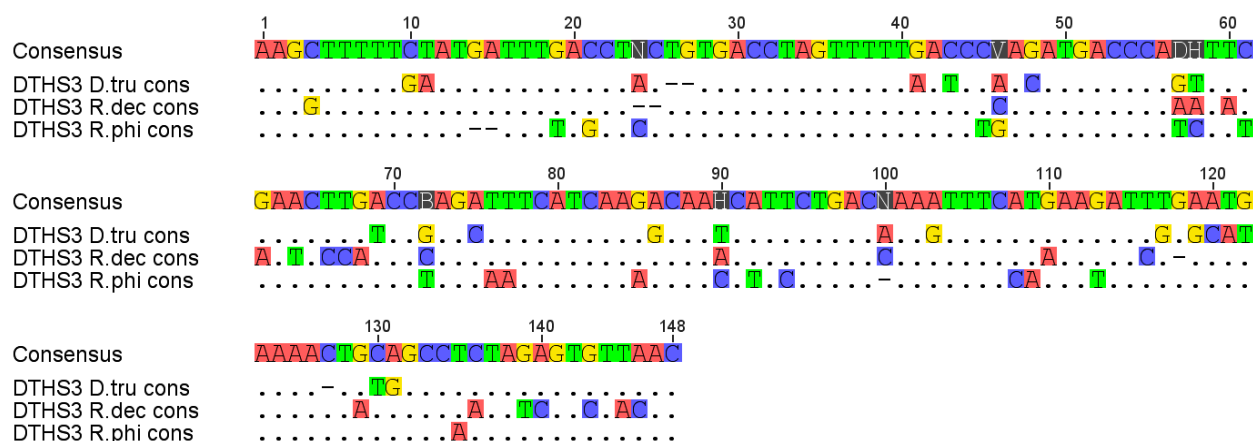


Slika 3.34: Svrnjenje monomera DTHS3 satelitne DNA izdvojenih iz klonova P46 i PD53, uz usporedbu s njihovom konsenzus sekvencom. Konsenzus sekvenca prikazana je punim nukleotidnim slijedom, a kod pojedinih jedinica ponavljanja prikazane su samo razlike u odnosu na konsenzus.

Satelitna DNA DTSH3 vrste *D. trunculus* poznata je od ranije (ref), a u vrstama *R. decussatus* i *R. philippinarum* uočena je tijekom ovog istraživanja. Ukoliko uspoređujemo nukleotidnu sličnost konsenzus sekvenci monomernih jedinica izdvojenih iz pojedinih vrsta (Tablica 3.7) uočavamo da je ona oko 70%. Svrnjenje konsenzus sekvenci satelitne DNA DTSH3 kod 3 vrste školjkaša prikazano je na Slici 35. Uočljivo je kako razina međusobne sličnosti nije podjednako raspoređena unutar jedinica ponavljanja, tj. postoji područje veće konzerviranosti u nukleotidnoj sekvenci, očuvano kod monomernih jedinica iz sve 3 vrste.

	<i>D. trunculus</i> kons.	<i>R. decussatus</i> kons.	<i>R. philippinarum</i> kons.
<i>D. trunculus</i> kons.	x	71,0%	73,0%
<i>R. decussatus</i> kons.	71,0%	x	72,6%
<i>R. philippinarum</i> kons.	73,0%	72,6%	x

Tablica 3.7: Stupanj sličnosti između nukleotidnih sljedova konsenzus sekvenci monomera satelitne DNA DTSH3 izolirane iz 3 vrste školjkaša.



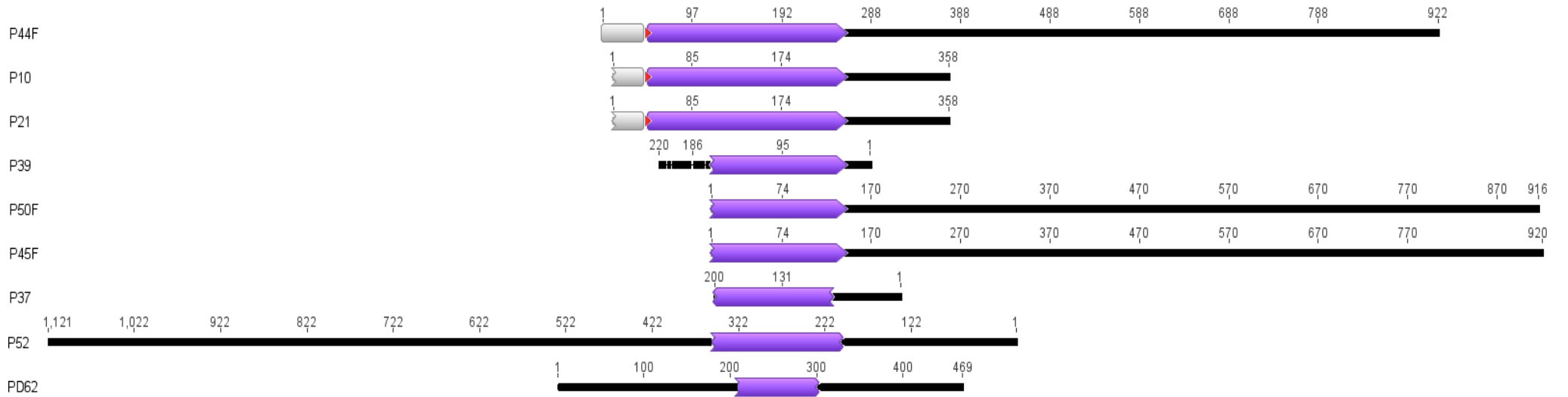
Slika 3.35: Svrnjenje konsenzus sekvenci satelitne DNA DTSH3 vrsta *D. trunculus*, *R. decussatus* i *R. philippinarum*. Njihova zajednička konsenzus sekvenca prikazana je punim nukleotidnim slijedom, a u konsenzus sekvencama pojedinih vrsta prikazane su samo nukleotidne razlike.

3.3.2. Element RPHI T1

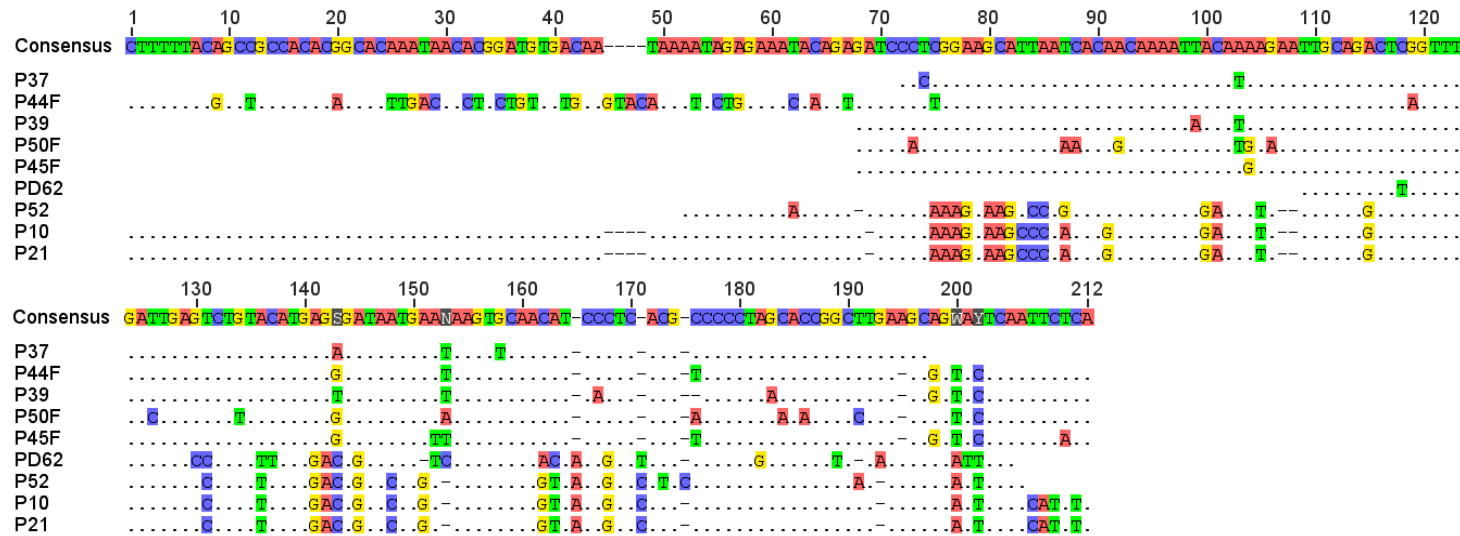
Koristeći sekvencu TDR elementa DTC T1 TRIM nađenog u vrsti *D. trunculus* provedena je *Local blast* pretraga unutar cjelokupne baze genomskih fragmenata školjkaša *R. philippinarum* te je uočena njegova prisutnost i kod ove vrste, u kojoj je, kao vrsno-specifična varijanta, dobio naziv RPHI T1. Element je pronađen kod sljedećih klonova i genomskih fragmenata: P10, P21, P37, P39, P44F, P45F, P50F, P52, PD62. U vrsti *R. philippinarum* nisu pronađeni elementi uobičajene strukture s dvije TDR domene, već varijante koje sadrže samo jednu TDR domenu (Slika 3.36). TDR domena ove vrste je duljine 205 pb. Broj deletiranih nukleotida kod krnjih varijanti je 47 - 104 na 5' kraju, odnosno 5 - 15 nukleotida na 3' kraju. Iz sravnjenja nukleotidnih sekvenci TDR domena uočljive su dvije grupacije elemenata RPHI T1 s obzirom na niz dijagnostičkih pozicija (Slika 3.37).

Unutarnja domena je kod ovih klonova prisutna samo u krnjim varijantama a nakon koje slijedi polipurinski trakt duljine 7 nukleotida, kao i kod vrste *D. trunculus*. Duplikacije na mjestu insercije nisu uočene, kao ni PBS mjesta. Također nije uočena nikakva konzerviranost u nukleotidnim sljedovima koji prethode elementima RPHI T1, kao ni kod onih koji slijede nakon njih u genomskom okolišu (izuzev kod klonova P10 i P21 koji dijele 100% homologije cijelom svojom dužinom).

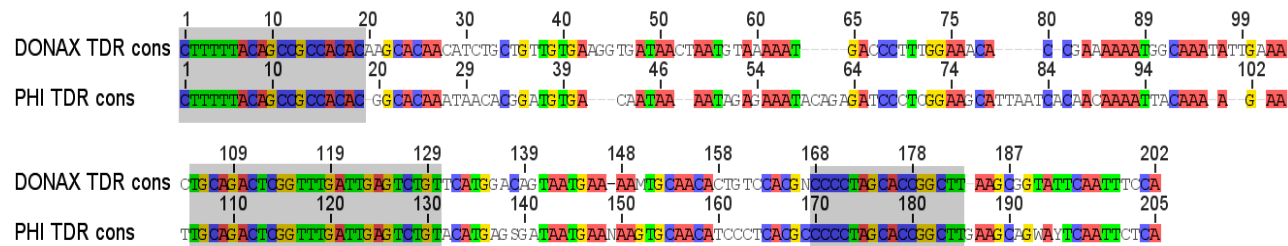
Sličnost između nukleotidnih sljedova konsenzus sekvenci TDR domena elemenata DTC T1 i RPHI T1 iznosi 71%, uz očito postojanje dobro konzerviranih regija unutar tih domena (Slika 3.38).



Slika 3.36: Genomski fragmenti koji sadrže elemente RPHI T1. TDR domene su označene ljubičasto, unutarnje domene sivo, polipurinski trakt crvenim strelicama.



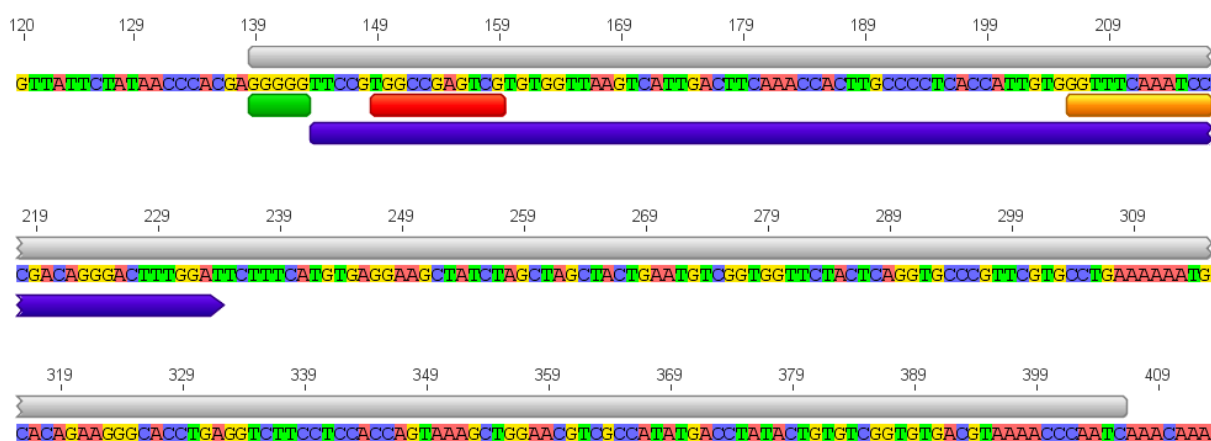
Slika 3.37: Svrnjenje nukleotidnih sljedova TDR domena elementa RPHI T1. Konsenzus sekvenca prikazana je punim nukleotidnim sljedom, a kod pojedinih jedinica ponavljanja prikazane samo razlike u odnosu na konsenzus.



Slika 3.38: Svrnjenje konsenzus sekvenci TDR domene elementa TRIM prisutnog kod vrsta *D. trunculus* i *R. philippinarum*. Konzervirani blokovi su sivo osjenčani.

3.3.3. Element RUDI 2

U sklopu genomskeg fragmenta P14F otkrivena je prisutnost SINE elementa istovjetnog s elementom RUDI 1, opisanim u vrsti *R. decussatus*. Element je nazvan RUDI 2, a nukleotidna sličnost među njima iznosi 80,9%. Na 5' kraju elementa koji je porijeklom od stanične tRNA, kod promotorskih sekvenci RNA polimeraze III, uočena je 1 mutacijska promjena u sklopu boxa A te jedna insercijska promjena u sklopu boxa B. Kratki G-bogati trakt na početku elementa je očuvan, dok poli A rep na 3' kraju elementa izostaje. Duplikacije na mjestu insercije također nisu uočene (Slika 3.39).



Slika 3.39: Nukleotidni sljed elementa RUDI 2 (sivo). Kratki G-bogati trakt označen je zeleno, regija porijeklom od tRNA ljubičasto, box A crveno, box B narančasto.

3.3.4. Klon P18

Genomski fragmenti vrste *R. philippinarum* izdvojeni u ovom radu (Tablica 3.5) su kategorizirani u skupine, pri čemu je navedeno da šesta skupina obuhvaća sve one koji sadrže više od jednog tipa repetitivne DNA te da se u njihovom sastavu mogu pronaći različite kombinacije elemenata prisutnih u svim prethodnim skupinama. Insert klona P18 (Slika 3.40) je odabran da bi se pobliže opisala ta skupina, kao i širok spektar različitih repetitivnih sekvenci koje se mogu naći unutar jednog genomske odsječka tog školjkaša. Duljina inserta klona P18 je 3995 pb. Područje od 130 – 204 pokazuje sličnost s non-LTR retrotranspozonom L1-56_XT, i to sa dijelom L1-56_XT2p koji kodira endonukleazu/reverznu transkriptazu (Kapitonov i Jurka, 2006). Područje od 770 - 1083 te od 1890 – 1992 pripada jednom monomeru satelitne DNA *phBgIII400* koji je nakon 312 nukleotida prekinut s 806 pb dugom insertiranom sekvencom, pri čemu je došlo do TGATC duplikacije na prekinutim krajevima monomera. Unutar 806 pb duge insertirane sekvence, *blastx* pretragom je uočena sličnost s reverznom transkriptazom mobilnog elementa *jockey* koji pripada elementima LINE. U nastavku monomera satelitne DNA *phBgIII400*, nakon kratke prijelazne sekvence od svega 14 nukleotida kojom jedna satelitna DNA postepeno prelazi u drugu, nalazi se 1427 pb dug niz koji pripada satelitnoj DNA BIV160. U zadnjem dijelu tog inserta, na području 3610 – 3688, uočena je sličnost u nukleotidnoj sekvenci s autonomnim DNA transpozonom EnSpm-5_VV.



Slika 3.40: Shematski prikaz inserta klona P18. Dio sekvence koji pokazuje sličnost sa non-LTR retrotranspozonom L1-56_XT je označen zeleno, fragmentirani monomer satelitne DNA *phBgIII400* je označen plavo, dio sekvence koji pokazuje sličnost s reverznom transkriptazom mobilnog elementa jockey ljubičasto, niz satelitne DNA BIV160 žuto, dio sekvence koji pokazuje sličnost s DNA transpozonom EnSpm-5_Vv crveno.

3.3.5. Usporedbe repetitivnog dijela genoma školjkaša *R. decussatus* i *R. philippinarum*

Kao što je već ranije spomenuto, da bi se utvrdilo postoji li sličnost i preklapanje u repetitivnim dijelovima genoma srodnih vrsta *R. decussatus* i *R. philippinarum* provedena je metoda otisaka kolonija uz korištenje genomske sonde jedne vrste na genomskoj biblioteci druge vrste. Kod školjkaša *R. decussatus* izdvojena su 4 klona koja su hibridizirala sa *R. philippinarum* genomskom sondom (Tablica 3.3). Među njima se može naći klon DP34 čiji insert jednim svojim dijelom pokazuje sličnost s elementom EnSpm11_SB. Dotični pripada porodici DNA transpozona EnSpm a usporedbom s Tablicom 3.6 možemo primjetiti da se radi o porodici koja je prisutna i kod vrste *R. philippinarum*. Klon DP36 koji sadrži satelitnu DNA DTHS3 daje hibridizacijski signal s genomskom sondom vrste *R. philippinarum* zahvaljujući prisutnosti te satelitne DNA u genomima obje vrste i više od 70% homologije među njima (Tablica 3.7). U analiziranom uzorku nije bio izdvojen niti jedan klon vrste *R. decussatus* koji bi hibridizirao s obje sonde.

Unutar genomske biblioteke vrste *R. philippinarum*, 19 klonova je hibridiziralo s *R. decussatus* genomskom sondom. Pritom je za 16 klonova (P14, P16, P18, PD22, P31, P43, P44, P45, P52, PD53, PD55, PD56, PD58, PD59, PD60, PD61) utvrđeno da daju hibridizacijski signal s obje genomske sonde. Među njima se ponovno nalazi jedan u čijem sastavu se nalazi satelitna DNA DTHS3 (klon PD53), kao posljedica već uočene sličnosti i prisutnosti u obje vrste. Još jedan klon koji je dao signal s obje hibridizacijske sonde, P14, sadrži element RUDI pronađen kod oba školjkaša, s ~80% homologije među njima. U sastavu nekih klonova koji daju signal s obje genomske sonde uočeni su različiti članovi porodice LTR transpozona Gypsy. No, zanimljivo je primjetiti da satelitna DNA *phBgIII400* zasad nije utvrđena u vrsti *R. decussatus*, ali prisutna je u čak nekoliko klonova koji daju hibridizacijski signal s genomskim sondama obje vrste (P16, P18, P43, PD60).

4. Rasprava

Značajan dio eukariotskog genetičkog materijala čine uzastopno i raspršeno ponovljene sekvence DNA. S obzirom da dominiraju brojem kopija u genomima raznih organizama, već su dugi niz godina predmetom istraživanja mnogih znanstvenih studija. Usprkos tomu, ponovljene sekvence DNA još uvijek predstavljaju nedovoljno poznatu genomsku komponentu nejasne evolucije i funkcionalnog značaja. S druge strane, školjkaši predstavljaju veliku i gospodarski značajnu skupinu organizama čiji su genomi općenito slabo istraženi pa se i sustavni genomski projekti ovih vrsta pojavljuju tek u najnovije vrijeme (Zhang i sur. 2012, Takeuchi i sur. 2012). Važno je istaknuti i da genomski projekti tek fragmentarno obuhvaćaju ponovljene sekvence DNA, prvenstveno zbog poteškoća u sklapanju dugih nizova međusobno vrlo sličnih jedinica ponavljanja pa su u tom pogledu nepotpuni (Eichler i sur. 2004). Stoga je središnja tema ovog rada istraživanje zastupljenosti i međusobnih odnosa ponovljenih sekvenci DNA i to u genomima 3 vrste školjkaša od gospodarskog i/ili ekološkog interesa: *Donax trunculus*, *Ruditapes decussatus* i *Ruditapes philippinarum*.

U sklopu ovog istraživanja detaljno je opisan novi raspršeno ponovljeni element, nađen u vrsti *D. trunculus* i nazvan DTC84 prema imenu klona u kojem je prvi put uočen. Ovaj element pokazuje modularnu organizaciju sličnu onoj skupine elemenata *pearl* (Gaffney i sur. 2003) i elementa *MgE* (Kourtidis i sur. 2007) prisutnima kod školjkaša, ali i niza mobilnih elemenata različitih vrsta kukaca, kao što su *DINE-1* (Yang i Barbash 2008), *SGM* (Miller i sur. 2000), *mini-me* (Wilder i Hollocher 2001), *PERI* (Kuhn i Heslop-Harrison 2011) te *MINE-2* (Coates i sur. 2011). Strukturne karakteristike koje objedinjuju ove elemente su dinukleotid AA kao TSD, centralna regija koja sadrži kratke nizove uzastopnih ponavljanja, prisutnost mikrosatelitne regije utemeljene na tetranukleotidnim ponavljanjima te subterminalno i terminalno smješteni obrnuto ponovljeni motivi. S obzirom na tetranukleotidni mikrosatelitni niz, iznimku predstavlja element *MgE* (Kourtidis i sur. 2007), kojemu nedostaje ova regija. Isti oblik strukturne organizacije uočen je kod još nekoliko elemenata otkrivenih tijekom ovog istraživanja, a koji istovremeno dijele i stanovitu nukleotidnu sličnost s nekim od elemenata skupine *pearl*. Tako dio genomskog fragmenta DTC12Alu vrste *D. trunculus* pokazuje sličnost s elementom CvE (Gaffney i sur. 2003), dijelovi fragmenta DTC37AluF i DTC41AluF dijele podjednaku sličnost s elementom CvG (Gaffney i sur. 2003) i s elementom DTC84, dok element nađen u sklopu klona D10 vrste *R. decussatus* pokazuje stanovit stupanj sličnosti s elementom CvA (Gaffney i sur. 2003). Elementi u fragmentima DTC37AluF i DTC41AluF pokazuju međusobnu sličnost od 90% te istovremeno iskazuju ~61% sličnosti s centralnim ponavljanjima elementa DTC84 i pripadajućim mikrosatelitnim

regijama. Ova osobina ih stavlja u kategoriju divergentnih varijanti DTC84 elemenata i objašnjava značajne razlike u udjelu tih sljedova u genomu školjkaša pri manje restriktivnim uvjetima hibridizacije, pri kojima je omogućena detekcija odvedenih varijanti.

Uzevši u obzir da se radi o neautonomnim mobilnim elementima bez kodirajućih regija, očekuje se dobra očuvanost samo obrnuto ponovljenih motiva na krajevima elementa, nužnih za transpoziciju (Craig 1995). Stoga nije čudno da broj centralnih ponavljanja unutar elemenata ove skupine može biti vrlo varijabilan, kao što je npr. kod elementa DINE (Yang and Barbash 2008), skupine *pearl* (Gaffney i sur. 2003) i elementa DTC84 opisanog u ovom radu. Na primjer, kod DTC84 su nađene varijante s 1–5 centralnih ponavljanja. Nukleotidni slijed odgovoran za varijabilan broj centralnih ponavljanja bi mogla biti palindromska sekvenca koja se nalazi pri kraju svake jedinice ponavljanja i koja je dobro sačuvana kod gotovo svih varijanti. Kod elementa DTC84 zanimljiv je i sam raspored varijanti centralnih ponavljanja. Uzevši u obzir varijabilnost u duljini niza i istovremeno dobro očuvanu organizaciju centralnih ponavljanja, teško je predvidjeti mehanizam koji bi sustavno održavao varijante 2* neposredno uz mikrosatelinu regiju, varijante grupe 2 nastavno na njih te varijante koje pripadaju grupi 1 na početku niza. Fiksna pozicija varijante 2* na kraju niza i prisutnost krnjih oblika na njegovom početku, ukazuju na mogućnost da su postojeći nizovi posljedica delecijских događaja izvornog niza, za kojeg bi se moglo pretpostaviti da je pentamer. Krnje varijante koje nastaju na početku niza mogle bi biti posljedica nedovoljno preciznih mehanizama koji dovode do rearanžmana, ali istovremeno i posljedica određenih preferencija tih mehanizama prema nukleotidnoj sekvenci, uzevši u obzir da postoji nekoliko pozicija u sklopu centralnog ponavljanja na kojima se prekid događa u većem broju slučajeva.

Unatoč tome što mogu varirati u broju kopija, centralna ponavljanja nađena unutar elemenata ovakvog tipa organizacije obično pokazuju visok stupanj homogenosti, odnosno varijabilnost njihove nukleotidne sekvence je vrlo mala (Gaffney i sur. 2003, Yang i Barbash 2008). Ovome se priklanja i element DTC84 kod kojeg je uočljivo >90% sličnosti među centralnim ponavljanjima.

Kao što je spomenuto, jedna od zajedničkih strukturnih karakteristika ovakvog tipa elementa je prisutnost mikrosatelitne regije. Bliska povezanost mikrosatelitnih nizova i raspršeno ponovljenih mobilnih genetičkih elemenata utvrđena je već ranije (Ramsay i sur. 1999, Wilder i Hollocher 2000). U prisutnosti takvih struktura, mikrosatelitna ponavljanja mogu se

svrstati u 3 kategorije: ona koja se nalaze na 3' kraju mobilnog elementa, ona koja se nalaze na 5' kraju te ona koja su nastala u unutrašnjosti elemenata. Prva kategorija je analogna ponavljanjima koja se mogu naći vezana uz SINE sekvence porodice Alu i nastaju kao posljedica integracije poliadeniliranog retro-transkripta u genom domaćina nakon čega slijede mutacije u poli-A nizu i produljenje u mikrosatelitni niz. Drugoj kategoriji pripadaju proto-mikrosatelitni nizovi (A bogati nizovi) koji su prisutni u genomu i poslužili su kao mjesto usidrenja mobilnog elementa. Treća kategorija obuhvaća one koji su se razvili kao sastavna komponenta mobilnog elementa, što bi bio slučaj kod DTC84. Wilder i Hollocher (2000) na temelju proučavanja mikrosatelitnih sekvenci *mini-me* elemenata prisutnih kod kukaca roda Diptera predlažu mehanizam nastanka ovakvih ponavljanja u dva koraka. U prvom koraku mutacijski događaji u sklopu početne nukleotidne sekvence dovode do nastanka dva uzastopna tetranukleotidna ponavljanja. U drugom koraku oni budu umnoženi tijekom replikacijskog proklizavanja, pri čemu dolazi do produljenja niza. U sklopu elemenata DTC84 opisanih u ovom istraživanju nađeni su nizovi tetranukleotidnih ponavljanja ACGG, koji mogu biti prekinuti tetranukleotidnom varijantom mutiranom u jednom ili dva nukleotida. Unatoč različitim duljinama nizova, po nekoliko ih dijeli zajedničke nukleotidne mutacije, što ukazuje na to da je sam mutacijski događaj nastao ranije u nekoj izvornoj kopiji, a do skraćivanja ili produljivanja niza došlo je naknadno. Dok se mikrosatelitna ponavljanja elementa CvE (Gaffney i sur. 2003) sastoje od motiva ACTG, kod njemu sličnog elementa u sklopu klona DTC12Alu mogu se naći malo promijenjene varijante te sekvence, pri čemu je 5 puta ponovljena varijanta ACCG i 4 puta ponovljena varijanta ACAG. Mikrosatelitna regija elementa koji se nalazi unutar klona DTC37Alu pokazuje još veću heterogenost jer sadrži 7 ponavljanja sastavljena od 3 različita motiva. Pri tomu se tetranukleotid ACGG (koji je prisutan kod njemu sličnih elemenata CvG (Gaffney i sur. 2003) i DTC84) nalazi u centralnom dijelu mikrosatelitne regije ponovljen 3 puta, a okružen je varijantom iste sekvence, ACTG. Na samom kraju niza može se naći i ACGA varijanta. U ovom slučaju je moguće da je došlo do mutacije u jednom tetranukleotidu usred ACTG niza, te njegove pretvorbe u ACGG i naknadnog umnažanja takve varijante unutar već postojećeg mikrosatelitnog niza.

Centralna ponavljanja 3 MITE elementa nađena kod vrste *C. virginica* koji pripadaju porodici *pearl* (CvA, CvE i CvG) pokazuju malo ili ništa međusobne nukleotidne sličnosti (Gaffney i sur. 2003), no, kao što je spomenuto, za svakoga od njih uočena je sličnost s nekim nukleotidnim sljedom prisutnim u drugim vrstama. Pritom, srodne varijante sva 3 elementa

prisutna u školjkašu *C. virginica*, također postoje i kod školjkaša *D. trunculus*. Isto tako, uočeno je paralelno postojanje divergentnih varijanti *DINE-I* elementa unutar 12 vrsta roda *Drosophila*, pri čemu je broj kopija pojedinih varijanti ovisan o pretpostavljenim naglim aktivacijama elemenata (Yang i Barbash 2008). Takve pojave ukazuju na činjenicu da bi i kod ovih raspršeno ponovljenih sekvenci mogla biti primjenjiva „teorija biblioteke“ koja opisuje evoluciju satelitnih DNA po kojoj se bilo koji član biblioteke može unutar pojedine vrste umnožiti i na taj način definirati vrsno specifični profil (Fry i Salser 1977, pregled u Plohl i sur. 2008).

U sklopu ovog istraživanja u vrsti *D. trunculus* je otkrivena i opisana vrsno-specifična varijanta satelitne DNA BIV160, široko rasprostranjene među školjkašima i pretpostavljene starosti oko 540 milijuna godina (Plohl i sur. 2010). Utvrđeno je da stupanj sličnosti između konsenzus sekvence Dt-BIV160 i konsenzus sekvence BIV160 satelitnih DNA prethodno utvrđenih kod nekoliko vrsta školjkaša iznosi 85,6%. S obzirom na postojeću homologiju, satelitna DNA Dt-BIV160 time može biti pridružena postojećoj skupini koja dijeli sličnost u nukleotidnoj sekvenci, a obuhvaća: satelitnu DNA *HindIII* prisutnu kod kamenica (Lopez-Flores i sur. 2004), satelitnu DNA DTE vrste *D. trunculus* (Plohl i Cornudella 1996), satelitnu DNA BIV160 nađenu kod različitih vrsta školjkaša (Plohl i sur. 2010) i mobilni element CvA prisutan kod kamenica (Gaffney i sur. 2003).

Osvrnuvši se i na samu duljinu jedinice ponavljanja ove nove satelitne DNA u školjkašu *D. trunculus*, može se uočiti da konsenzus duljina monomera Dt-BIV160 iznosi 166 pb i jednaka je konsenzus duljini jedinica ponavljanja skupine BIV160 prisutne kod vrsta porodice Veneridae (166 ± 1 pb, Plohl i sur. 2010) te konsenzus duljini jedinice ponavljanja centromerne satelitne DNA *HindIII* u kamenicama (166 ± 2 pb, Clabby i sur. 1996, Lopes-Flores i sur. 2004), a slična je onima kod većine dosad otkrivenih satelitnih DNA u razredu Bivalvia (Plohl i Cornudella 1996, Plohl i Cornudella 1997, Martinez-Lage i sur. 2002, Canapa i sur. 2000, Petrović i Plohl 2005). Međutim, u školjkašima su otkrivene i satelitne DNA duljine monomera 40 pb (Biscotti i sur. 2007), te one duljine 400 pb (Passamonti i sur. 1998), pri čemu se u ovom svojstvu satelitne DNA školjakaša ne razlikuju od satelitnih DNA općenito. Iako duljine jedinica ponavljanja satelitnih DNA mogu varirati u rasponu od nekoliko desetaka parova baza do više kilobaza, najčešće se kreću u rasponu od 100-400 pb. Duljina monomera većine centromernih satelita nalazi se pak u rasponu od 150 - 210 pb, što odgovara duljini nukleosomskih jedinica (Henikoff i sur. 2001). U tom kontekstu zanimljivo je

primijetiti i da su uzastopna ponavljanja smještena unutar raspršeno ponovljenih elementa pearl ~160 pb (Gaffney i sur. 2003), što pak odgovara duljini jedinice ponavljanja većine satelitnih DNA ovih školjkaša.

Usporedba nukleotidnih sljedova konsenzus sekvenci satelitnih DNA skupine BIV160, *HindIII* i DTE te centralnih ponavljanja mobilnog elementa CvA pokazala je kako nukleotidna varijabilnost nije podjednako raspoređena unutar jedinica ponavljanja. Drugim riječima, u jedinici ponavljanja postoje nukleotidni segmenti koji su više sačuvani i oni u kojima je razlika u sekvenci izraženija. Postojanje područja veće konzerviranosti u nukleotidnoj sekvenci je zanimljivo, uzevši u obzir da se radi o sekvencama izdvojenim iz udaljenih vrsta školjkaša. Zatim, kod satelitne DNA DTHS3 koja je u vrsti *D. trunculus* poznata od ranije (Plohl i Cornudella 1997), a u vrstama *R. decussatus* i *R. philippinarum* je otkrivena tijekom ovog istraživanja, nakon usporedbe konsenzus jedinica ponavljanja također je vidljivo područje visoke konzerviranosti u dijelu nukleotidne sekvence, unatoč tome što se radi o jasno prepoznatljivim vrsno-specifičnim varijantama. Nejednoliki raspored mutacija unutar sekvenci monomera zapažen je kod satelitnih DNA prisutnih i u raznim drugim vrstama, npr. kod α -satelitne DNA čovjeka, kod satelitnih DNA vrste *Arabidopsis thaliana* (Hall i sur. 2003) i kukaca roda *Tribolium* (Mravinac i sur. 2004), kod DTF1 satelitne DNA školjkaša *D. trunculus* (Petrović i Plohl 2005), te u nematoda (Meštrović i sur. 2006). Izmjenjivanje varijabilnih i sačuvanih područja u nukleotidnoj sekvenci satelitnih monomera dovodi se u vezu s funkcionalnim pritiscima, odnosno s mogućim interakcijama s proteinima koji vežu DNA i koji bi mogli biti uključeni u izgradnju heterokromatina (Hall i sur. 2003). Dobro poznati motiv sekvence DNA uključen u interakciju s centromernim proteinima je 17 pb dug slijed CENP-B, nađen unutar dijela varijanti monomera α -satelitne DNA čovjeka. Taj motiv prepoznaje protein CENP-B (Masumoto i sur. 1989) i pretpostavlja se da sudjeluje u formiranju centromernog heterokromatina (Ohzeki i sur. 2002). Kod sisavaca su CENP-B motiv i pripadni protein visoko konzervirani, a varijante tog motiva nađene su kod beskralješnjaka, uključujući i školjkaše (Canapa i sur. 2000). Uočena je značajna sličnost između proteina CENP-B i transpozaze mobilnih elemenata *pogo* (Tudor i sur. 1992), što bi ukazivalo na njegovo porijeklo od nadporodice transpozaza *Tc1/mariner*. Usporedba aminokiselinskog slijeda CENP-B i njemu srodnih proteina otvara mogućnost da CENP-B također posjeduje mogućnost stvaranja rezova u jednom lancu molekule DNA. Kipling i Warburton (1997) predlažu da bi takva mjesta u blizini motiva CENP-B promovirala rekombinaciju i imala važnu ulogu u rearanžmanima satelitnih DNA.

Vezano uz genetičke rearanžmane, u prilog ideji da bi centralna ponavljanja kao i monomerne jedinice satelitnih DNA mogle funkcionirati kao zasebne insercijske i delecijske jedinice ide pronalazak monomera Dt-BIV160 u blizini elementa DTC84 u jednom od genomskih klonova (84-35) izdvojenih tijekom ovog istraživanja. Uzevši u obzir da je BIV160 sekvenca srodna centralnim ponavljanjima elementa CvA, možemo zaključiti da je taj nukleotidni slijed prisutan u 3 različita organizacijska oblika: kao centralna ponavljanja unutar mobilnih elemenata, kao uzastopno ponovljeni monomeri tvoreći nizove satelitne DNA (Plohl et al. 2010) te kao zasebne kopije raspršene u genomu. Treći organizacijski oblik je potvrđen cijelim nizom klonova izdvojenih tijekom ovog istraživanja, kod kojih se mogu naći samo kraći fragmenti satelitne DNA vezani uz neke druge genomske sljedove. Tako kod vrste *D. trunculus* DTC17AluF i DTC51Alu sadrže fragmente satelitne DNA Dt-BIV160, a DTC38Alu fragment satelitne DNA DTE. Kod vrste *R. decussatus* u sklopu D7F, D26 i DP36 se mogu naći fragmenti koji pripadaju satelitnoj DNA DTHS3, dok su kod vrste *R. philippinarum* kod P12, P13, P15, P16, P34, P35, P41, P42, P43, P46 i PD60 u pitanju fragmenti satelitne DNA *phBgIII400*. Monomerne jedinice ili kratki nizovi od 1-5 jedinica primijećeni su i u eukromatinskim područjima vrste *Drosophila melanogaster* (Kuhn i sur. 2012), u blizini gena ili u sklopu introna, a isti slučaj je zamijećen i kod kukca kornjaša *Tribolium castaneum* (Brajković i sur. 2012). Kod kukca *T. castaneum* je utvrđeno da se sekvenca satelitne DNA TCAST pojavljuje u različitim organizacijskim oblicima: dugim nizovima uzastopno ponovljenih monomera, u obliku parcijalnih monomernih jedinica ili kratkih nizova (1 - 4 jedinice), u strukturama nalik DNA transpozonu, te u 5' UTR regiji retrotranspozona CR1-3 Tca. Posljednje dvije strukture objašnjavaju mobilnost ove sekvence i činjenicu da se može pronaći i u eukromatinskom dijelu genoma, u neposrednoj blizini gena. Isto tako, uzastopno ponovljena sekvenca koja odgovara centralnim ponavljanjima elementa CvA školjkaša *C. gigas* je pronađena unutar 5' UTR regije pretpostavljenog gena za transglutaminazu (Gaffney i sur. 2003).

Iako nastanak početnih nizova satelitnih DNA još uvijek nije razjašnjen, sličnost u sekvenci s mobilnim elementima u genomu dovodi do pretpostavke da su oni izvor jedinica ponavljanja satelitnih DNA. Već je spomenuta sličnost između MITE mobilnog elementa CvA i cijelog niza satelitnih DNA (Dt-BIV160, ovaj rad; BIV160, Plohl i sur. 2010; *HindIII*, Lopez-Flores i sur. 2004; DTE, Plohl i Cornudella 1996). U sklopu ovog istraživanja uočena je poveznica između još dvije porodice satelitnih DNA i mobilnih elemenata. Jedna od njih je DTHS1 (Plohl i Cornudella 1997), a pokazuju sličnost s početnim dijelom non-LTR retrotranspozona

Coprina_Cc1, koji je uglavnom vezan za telomerne sekvence (Gladyshev i Arkhipova 2007). Druga porodica je DTHS2 (Plohl i Cornudella 1997), koja cijelom svojom dužinom monomera pokazuje relativno niski stupanj homologije (62%) s dijelom LTR retrotranspozona Gypsy-144_AA-I (Nene i sur. 2007). Značajna sličnost uočena je u sekvenci satelitne DNA pvB370 i LTR sekvenci transpozona pDv skupine vrsta *Drosophila virilis* (Heikkinen i sur. 1995). Sličan odnos primijećen je i za satelitnu DNA Sobo prisutnu kod krumpira (Tek i sur. 2005), a poznati su i brojni drugi primjeri sličnosti satelitnih DNA i mobilinih elemenata, odnosno njihovih dijelova (Li i Leung 2006, Kapitonov i sur. 1998, Kapitonov i Jurka 1999, Miller i sur. 2000). U genomima vrsta roda *Xenopus* nađene su dvije uzastopno ponovljene sekvence, Xstir i tropXstir, obje nalik različitim dijelovima MITE elementa Xmix (također prisutnom u tim vrstama), na osnovu kojih je predložen princip nastanka satelitnih DNA iz MITE elemenata (Hikosaka i Kawahara 2004). Prema ovom modelu, nastanak uzastopnih ponavljanja unutar samog MITE elementa uključuje stvaranje ukosnica između dva susjedna MITE elementa na jednolančanoj DNA (posljedica kašnjenja jednog lanca tijekom replikacije DNA) te izrezivanje dijela ukosnice djelovanjem nukleaza. Kao posljedica ovog procesa, novi MITE element će sadržavati dva puta uzastopno ponovljeni dio sekvence, a istim procesom može doći i do daljnjeg produljenja niza. Takvi kratki nizovi još uvijek predstavljaju raspršeno ponovljene sekvence, no uz sudjelovanje rekombinacijskih procesa oni bi se mogli razviti u dugačke uzastopno ponovljene nizove, odnosno u nove satelitne DNA. Drugi mehanizam nastanka uzastopnih ponavljanja također se temelji na formiranju ukosnica (Izsvak i sur. 1999). Tijekom replikacije MITE elementa, zahvaljujući invertnim ponavljanjima ili palindromskim sekvencama, došlo bi do stvaranja ukosnice u novonastalom lancu koji je još u procesu sinteze, nakon čega dolazi do izvrtnja cijele strukture na način da ona sama sebi postaje kalup. Kada replikacija takve strukture dođe do kraja, duplicirani segment se odvaja s mjesta sinteze i ugrađuje se na novo mjesto u genomu. Međutim, sigurno postoje i drugi mehanizmi odnosno putevi koji dovode do istih posljedica. U ovom trenutku možemo tek nagađati da se ovdje radi o cijelom nizu procesa i događaja vezanih uz replikaciju i rekombinaciju, a preferenciju prema pojedinima mogle bi određivati specifičnosti u samoj nukleotidnoj sekvenci raspršeno ponovljenih elemenata.

Osim kratkih i neautonomnih varijanti DNA transpozona (MITE elementi), u sklopu ovog istraživanja uočeni su kratki i neautonomni LTR retrotranspozoni (TRIM elementi). Njihova veličina (<540 pb) je značajno manja od one tipične za LTR retrotranspozone (5–15 kb), što ih čini najkraćim mobilnim elementima s TDR sljedovima otkrivenim do sada. Ovakav tip

elemenata je nađen kod biljaka (Witte i sur. 2001, Yang i sur. 2007) i kod različitih porodica mrava (Zhou i Cahan 2012), a u sklopu ovog istraživanja po prvi put je uočen i unutar razreda Bivalvia. Kod mrava *Pogonomyrmex barbatus* (Zhou i Cahan 2012), od 270 izvojenih kopija elementa PbTRIM, svega 67 je bilo cjelovito, dok su prevladavale varijante samo s jednom TDR domenom (159 elemenata) i krnje varijante (44 elementa). Slično tome, svi primjerci TRIM elementa DTC T1 iz vrste *D. trunculus* su krnje varijante, pri čemu je nađena jedna kopija s 3 TDR domene i jedna kopija koja sadrži samo jednu TDR domenu. Srodni element RPHI T1 iz vrste *R. philippinarum* nađen je samo u oblicima s jednom TDR domenom, redovito krnjom. Pretpostavlja se da su takvi primjerci s jednom TDR domenom posljedica događaja nejednolikog prekrivanja (*crossing-over*) između TDR domena TRIM elemenata (Bennetzen i Kellog 1997), a istim procesom bi mogle nastati i varijante s 3 TDR domene. Kod TRIM elemenata iz biljaka (Witte i sur. 2001), kao i kod elemenata PbTRIM otkrivenih kod mrava *Pogonomyrmex barbatus* (Zhou i Cahan 2012), postoji niz karakteristika koje ukazuju na njihovu nedavnu aktivnost. To uključuje: visok stupanj homologije među TDR domenama pojedinog elementa, veliku međusobnu sličnost kopija s različitih genomskih lokacija, te prisutnost TSD motiva kod gotovo svih kompletnih elemenata (neovisno o broju TDR domena). Sve te karakteristike koje bi ukazivale na nedavnu aktivnost izostaju kod elemenata DTC T1 i RPHI T1, izdvojenih u sklopu ovog istraživanja. Kod elemenata DTC T1 središnji dio elementa je bolje očuvan, dok TDR sekvence iskazuju stupanj homologije od 71,9% - 92,7% i redovito su pretrpile ponekad i znatne, delecije rubnih dijelova. Također nedostaju TSD mjesta kao posljedica insercije. Kod elementa RPHI T1 je situacija gotovo identična, iako je nađen samo u varijantama s jednom TDR domenom.

Cjelovite varijante elementa PbTRIM nađene kod mrava *Pogonomyrmex barbatus* (Zhou i Cahan 2012), grupiraju se u 4 potporodice. Početnicama specifičnim za konzervirane dijelove TDR domena bili su izdvojeni elementi iz još 8 vrsta mrava, pri čemu su svi oni koji su bili porijeklom izvan vrste *Pogonomyrmex* pripadali potporodici 1. Odnosi među elementima u toj potporodici ukazivali su na potencijalni horizontalni prijenos u širenju elemenata PbTRIM, s obzirom na vrlo visoku sličnost u nukleotidnoj sekvenci među elementima izdvojenih iz međusobno udaljenih vrsta. Tako je, npr. PbTRIM iz *Pogonomyrmex* J-linije bio 100% identičan onome iz vrste *Pheidole desertorum*, a PbTRIM elementi iz vrsta *L. neoniger* i *P. barbatus*, koje pripadaju različitim potporodicama razdvojenim prije više od 100 milijuna godina, pokazuju 99% međusobne sličnosti. Usporedbom elemenata TRIM izdvojenih u sklopu ovog istraživanja, nema nikakvih naznaka da bi horizontalni prijenos bio prisutan

između vrsta *D. tunculus* i *R. philippinarum*. Sličnost između nukleotidnih sljedova konsenzus sekvenci TDR domena elementa DTC T1 i RPHI T1 iznosi 71%, no nema nikakvih naznaka njihove nedavne mobilnosti. Oni vrlo vjerojatno potječu od elementa koji je bio prisutan u zajedničkom pretku, nakon čega su se razvile vrsno-specifične varijante. Uzevši u obzir vrijeme nastanka reda Veneroida (Plohl i sur. 2010), pretpostavljena starost ovih elementa bila bi oko 250 milijuna godina. Za pretpostaviti je da je i kod vrste *R. philippinarum* izvorno bila prisutna verzija s 2 TDR domene, a nakon rekombinacijskog događaja uslijed kojeg je nastala RPHI T1 varijanta s jednom TDR domenom došlo je do umnažanja novonastale varijante u većem broju kopija. Postojanje varijanti s više TDR domena u toj vrsti nije isključeno, iako nisu bile uočene u sklopu ovog istraživanja.

Treba spomenuti da genomi različitih vrsta često pokazuju različitu preferenciju prema nakupljanju jednog ili drugog tipa raspršeno ponovljenih sekvenci u genomu. Tako su kod kvasca *Saccharomyces cerevisiae* prisutni isključivo retrotranspozoni, dok je dominacija DNA transpozona uočena kod prazivotinje *Entamoeba invadens* (Feschotte i Pritham 2007). Kod školjkaša *D. trunculus*, *R. decussatus* i *R. philippinarum* takva preferencija nije uočena, uzevši u obzir podatke iz Tablice 3.2, 3.4 i 3.6 te gledajući relativnu zastupljenost elemenata. Oni pokazuju laganu dominaciju transpozona u odnosu na retrotranspozone, osim vrste *R. decussatus* u kojoj je zastupljenost DNA transpozona nešto znatnija.

Rezultat analize nukleotidnih sljedova genoma školjkaša *Pinctada fucata* (Takeuchi i sur. 2012) veličine ~1150 Mb, pokazuje da 9,8% genoma ove vrste obuhvaćaju ponavljajuće sekvence. Pri tome 0,4% čine DNA transpozoni, 1,5% retrotranspozoni i 7,9% uzastopno ponovljene sekvence DNA, od čega 1,85% pripada satelitnim DNA. Sekvence LINE i SINE zauzimaju iznimno male postotke, 0,3 i 0,08%. U ovom slučaju je relativna zastupljenost raspršeno ponovljenih sekvenci na strani retrotranspozona. Odrediti koji udio čine raspršeno ponovljene sekvence unutar dijela genoma kojem je utvrđen nukleotidni slijed u sklopu ovog istraživanja te ih usporediti s udjelima utvrđenim za vrstu *Pinctada fucata* nije moguće, iako bi takva usporedba bila informativna. Razlog tome je da metoda kojom su izdvojeni genomske fragmenti u ovom istraživanju preferira ponavljajuće sekvence, pa je time puno veća njihova relativna zastupljenost u analiziranom dijelu genoma. Međutim, porodice DNA transpozona nađene u vrsti *Pinctada fucata* (Mariner, Polinton, Helitron, Academ, PIF-Harbinger, P, Zator, Sola), kao i porodice LTR retrotranspozona (Gypsy, DIRS, and Bel_Pao) te non-LTR

retrotranspozona (Penelope i CR1) se većinom preklapaju s onima nađenima u vrstama *D. trunculus* (Tablica 3.2), *R. decussatus* (Tablica 3.4) i *R. philippinarum* (Tablica 3.6).

Nakon utvrđivanja nukleotidnih sljedova genoma još jednog školjkaša, *C. gigas*, veličine 637 Mb (Zhang i sur. 2012), utvrđeno je da ponavljajuće sekvence čine 36% genoma, što je gotovo četiri puta više nego kod vrste *P. fucata* (Takeuchi i sur. 2012). Ustanovljeno je da DNA transpozoni čine 4,1%, LTR retrotranspozoni 2,5%, LINE 1,4%, SINE 0,6%, uzastopno ponovljena DNA 3,6%, a nepoznate repetitivne sekvence čak 22,6% genoma. Uočen je velik broj gena za transpozazu (359) i reverznu-transkriptazu (779), kao i prisutnost njihovih transkripata. Također, nađeno je 157 000 kopija MITE elemenata, koji čine znatnih 8,82% genoma. Sravnjenja nukleotidnih sekvenci su pokazala iznimno malu varijabilnost unutar nekih porodica MITE elemenata, ukazujući na mogućnost da su ti elementi još uvijek aktivni. Kod ovog školjkaša također postoji preferencija prema DNA transpozonomima, pogotovo ako uzmemo u obzir velik broj nađenih MITE elemenata. U vrsti *D. trunculus* nađene su 3 porodice MITE elementa kod kojih je očuvanost u nukleotidnoj sekvenci prilično visoka: 92,6% za centralna ponavljanja elementa DTC84 te u rasponu od 85,1% - 100% za cjelovite elemente DTC M1, odnosno 83,2% - 100% za DTC M2. To bi ih svrstalo u potencijalno još uvijek aktivne mobilne elemente, iako je u genomu prisutan i velik broj krnjih varijanti DTC M1 i DTC M2, kao i odvedenih varijanti DTC84. Možemo nagađati da i kod vrste *D. trunculus* značajan dio genoma obuhvaćaju DNA transpozoni, uzevši u obzir da element DTC84 sam čini 1% genoma, a ukoliko uzmemo u obzir i odvedene varijante, njegova zastupljenost je procijenjena do 2%.

Prilikom utvrđivanja postoji li sličnost i preklapanje u repetitivnim dijelovima genoma srodnih vrsta *R. decussatus* i *R. philippinarum* analizirani su hibridizacijski otisci kolonija s kloniranom genomskom DNA uz korištenje genomske sonde jedne vrste na genomskoj biblioteci druge vrste. Kod školjkaša *R. decussatus* izdvojena su 4 klona koja hibridiziraju s genomskom sondom vrste *R. philippinarum* (Tablica 3.3). Među njima se može naći klon čiji insert jednim svojim dijelom pokazuje sličnost s elementom EnSpm11_SB, a dotični pripada porodici DNA transpozona EnSpm koja je prisutna i kod vrste *R. philippinarum*. Jedan od klonova izdvojenih tom metodom sadrži satelitnu DNA DTHS3, a možemo zaključiti da pokazuje hibridizacijski signal s genomskom sondom vrste *R. philippinarum* zahvaljujući prisutnosti te satelitne DNA u genomima obje vrste, s više od 70% homologije među njima (Tablica 3.7). Prisutnost te satelitne DNA kod obje vrste je po prvi put utvrđena u sklopu

ovog istraživanja. Unutar genomske biblioteke vrste *R. philippinarum*, 19 klonova je hibridiziralo s genomskom sondom vrste *R. decussatus*. Za 16 klonova je utvrđeno da daju hibridizacijski signal s obje genomske sonde. Među njima se ponovno nalazi jedan klon u čijem sastavu je satelitna DNA DTSH3, kao posljedica već uočene sličnosti i prisutnosti u obje vrste. Još jedan klon koji je dao signal s obje hibridizacijske sonde sadrži SINE element RUDI, pronađen kod oba školjkaša, s ~80% homologije među njima. U sastavu još nekih klonova vrste *R. philippinarum*, koji daju signal s obje genomske sonde, uočeni su različiti članovi porodice LTR transpozona Gypsy zahvaljujući prisutnosti te porodice kod obje vrste. Zanimljivo je primjetiti da satelitna DNA *phBgIII400* zasad nije bila utvrđena u vrsti *R. decussatus* (Passamonti i sur. 1998), ali prisutna je u čak nekoliko klonova koji daju hibridizacijski signal s genomskim sondama obje vrste. Ovi bi rezultati ukazivali na veliku vjerojatnost njezine prisutnosti i kod tog školjkaša, a neobjavljeni rezultati (Mravinac 2013) donose i konačnu potvrdu njezine prisutnosti. Uzevši u obzir informacije dostupne do sad, možemo zaključiti da srodne vrste *R. decussatus* i *R. philippinarum* (Veneridae) dijele 3 zajedničke porodice satelitnih DNA: BIV160 (Plohl i sur. 2010), DTSH3 (ovo istraživanje) i *phBgIII400* (Passamonti i sur. 1998, ovo istraživanje, Mravinac 2013). Pri tomu, dvije porodice satelitnih DNA BIV160 i DTSH3 dijele i s udaljenijom vrstom *D. trunculus* (Donacidae). Satelitna DNA BIV160 je pretpostavljene starosti od oko 540 milijuna godina (Plohl i sur. 2010), dok je starost DTSH3 teško odrediti jer do sad nije nađena izvan reda Veneroida kojem pripadaju sve 3 istražene vrste. Stoga je njezina pretpostavljena minimalna starost 250 milijuna godina. Za satelitnu DNA *phBgIII400* možemo pretpostaviti da se pojavila tek kasnije, jer je za sad utvrđena samo u sklopu porodice Veneridae, kod vrsta *R. philippinarum*, *Paphia undulata*, *Venerupis aurea* (Passamonti i sur. 1998) i vrste *R. decussatus* (Mravinac, neobjavljeni rezultati).

Yoshida i sur. (2011) su proveli istraživanje na 3 vrste unutar koljena mekušaca, želeći utvrditi koliki utjecaj imaju ponovljene sekvence DNA na veličinu genoma. U istraživanje su bili uključeni glavonošci *Idiosepius paradoxus* i *Nautilus pompilius* te školjkaš *Mizuhopecten yessoensis*. Utvrđeno je da su veličine genoma tih vrsta: 2,1 Gb za *I. paradoxus*, 4,2 Gb *N. pompilius* i 1,8 Gb *M. yessoensis*, pri čemu udio ponovljenih sekvenci u genomima iznosi, istim redom, 9,2%, 4,0%, i 3,8%. Usporedbom dobivenih rezultata autori zaključuju da širenje ponovljenih sekvenci nije značajno utjecalo na veličinu genoma vrste *N. pompilius*. U usporedbi s glavonošcem *I. paradoxus*, ova vrsta ima duplo veći genom, ali značajno manji udio ponovljenih sljedova, što bi ukazivalo na to da su povećanju genoma u ovom slučaju

doprinjeli neki drugi faktori mimo ponovljenih sekvenci DNA. Također, udio retrotranspozona u genomima ta dva glavonošca je značajno veći kod vrste *I. paradoxus*. Kod glavonožaca je prethodno utvrđena jedna porodica SINE sekvenci, karakteristična baš za tu skupinu, Ceph-SINE (Akasaki et al., 2010). Yoshida i sur. (2011) su uočili da se kod tih sekvenci iza Ceph domene nalazi regija porijeklom od 7S RNA iz genoma školjkaša, što bi ukazivalo na to da Ceph-SINE zapravo potječe od zajedničkog pretka glavonožaca i školjkaša (Sigwart and Sutton, 2007). Kod glavonošca *N. pompilius* je ta retrosekvenca prisutna u manjem broju nego kod glavonošca *I. paradoxus* i školjkaša *M. yessonensis*.

U sklopu ovog istraživanja nađena je samo jedna familija SINE sekvenci kod vrste *R. decussatus* (RUDI 1) te njoj srodna RUDI 2 u vrsti *R. philippinarum*. Obje su zastupljene u malom broju klonova, a Ceph-SINE ovim putem nije utvrđena kod tih vrsta. Element RUDI 1 ima očuvane TSD sekvence kao i motive box A i box B i vjerojatno je aktivan, dok kod RUDI 2 elementa TSD sekvence izostaju, a box A i box B su pretrpili po jednu mutacijsku, odnosno insercijsku promjenu.

U sklopu istraživanja Yoshida i sur. (2011) određen je i broj kodirajućih gena kod 3 mekušca, on iznosi 22 300 za *I. paradoxus*, 39 080 za *N. pompilius* i 34 180 za *M. yessonensis*. Procjena broja kodirajućih gena omogućuje i usporedbu pojedinih genskih porodica među tim organizmima. Porodica kojoj pripada najveći broj gena kod vrste *I. paradoxus* su geni za „cink-finger“ proteine i obuhvaća 154 različita gena prisutna u više od 3000 kopija u genomu. Kod vrsta *N. pompilius* i *M. yessonensis* nađeno je svega 5, odnosno 3 takva gena. Pretpostavlja se da je njihov veliki broj kod vrste *I. paradoxus* posljedica duplikacija gena, poput onih uočenih kod Hamilton i sur. (2003) te Tang i sur. (2002). Određen broj klonova vrsta *D. trunculus*, *R. decussatus* i *R. philippinarum* pokazuje sličnost s nukleotidnim slijedovima koji se nalaze u blizini gena za „cink-finger“ protein školjkaša *Spisula solidissima* i gena Msx (jedan od gena regulatora ontogenije kod Metazoa) školjkaša *Corbicula fluminea*, no, s obzirom da se u većem broju slučajeva ne radi o samim kodirajućim sekvencama, za pretpostaviti je da su sličnost pokazale neke za sad neidentificirane raspršeno ponovljene sekvence prisutne u relativnoj blizini tih gena, a koje su se tamo našle nakon insercijskih događaja ili kao posljedica genetičkih rearanžmana.

Najbrojnija porodica gena školjkaša *M. yessonensis* (Yoshida i sur. 2011) je ona za endonukleazu/reverznu-transkriptazu. Retrotranspozonskog je porijekla i obuhvaća 75 gena

prisutnih u 1200 kopija, dok kod glavonožaca *N. pompilius* i *M. yessonensis* homolozi tih gena nisu izdvojeni. Slično tome, u sklopu 3 klonova vrste *D. trunculus* pronađeni su nukleotidni sljedovi koji pokazuju sličnost s genom za reverznu transkriptazu (Tablica 3.1). Kod vrste *R. decussatus* nađen je jedan klon koji sadrži slijed za reverznu transkriptazu i jedan za replikazu/helikazu/endonukleazu (Tablica 3.3), a kod vrste *R. philippinarum* tri klonova pokazuju sličnost s genom za reverznu transkriptazu/endonukleazu (Tablica 3.5).

Sve 3 vrste školjkaša, čiji su repetitivni dijelovi genoma bili predmetom ovog istraživanja, sadrže jednu grupu klonova u kojoj se nalaze nukleotidni sljedovi kod kojih je uočena istovremena prisutnost različitih vrsta ponovljenih sekvenci unutar kratkih genomskih odsječaka. Insert klona P18 vrste *R. philippinarum* je odabran da bi se pobliže opisala ta skupina, kao i širok spektar različitih ponovljenih sekvenci koje se mogu naći unutar jednog genomskog odsječaka školjkaša, kao i njihova međusobna isprepletenost u izgradnji „mreže“ ponovljenih sljedova. U ovom slučaju, dvije različite porodice satelitnih DNA nalaze se u neposrednoj blizini, gotovo nastavno jedna na drugu. Pritom se satelitna DNA BIV160 nalazi u obliku niza uzastopno ponovljenih monomera, dok je *phBgIII400* prisutna u obliku samo jednog monomera, prekinutog insercijom raspršeno ponovljenog elementa uz nastanak pentanukleotidnog TSD mjesta. U sklopu elementa uočena je sličnost s reverznom transkriptazom, što bi značilo da se radi o autonomnom retrotransponirajućem elementu. Nedaleko monomera satelitne DNA *phBgIII400* nalazi se nukleotidni slijed koji pokazuje sličnost s dijelom non-LTR retrotranspozona, a nedaleko niza satelitne DNA BIV160 slijed koji pokazuje sličnost s dijelom DNA transpozona. Slična situacija, insercija mobilnog elementa u monomer satelitne DNA uočena je i u sklopu klona D12 vrste *R. decussatus*, u kojem se element RUDI 1 nalazi unutar monomera satelitne DNA DTHS3. Osim kod velikog broja klonova u ovom istraživanju, prisutnost različitih rearanžmana i kombinacija koje uključuju uzastopno i raspršeno ponovljene sljedove DNA uočena je i u sklopu istraživanja kojeg su Palomeque i sur. (2006) proveli na kukcima roda *Messor*. Oni donose informaciju o inserciji raspršeno ponovljenih elemenata IRE130 u nizove satelitne DNA. IRE130 je neautonomni MITE duljine 130 pb i cijelom svojom duljinom je nesavršena palindromska sekvenca. Osim što se uvijek preferencijalno ugrađuje na isto mjesto unutar monomera satelitne DNA, on se istovremeno i sam pokazao kao dobro ciljno mjesto za daljnje ugradnje. Naime, nađeno je nekoliko primjera kada se u sklopu elementa IRE130 nalazi ugrađen nepoznati raspršeno ponovljeni slijed duljine 851 pb, kao i nekoliko primjera kada je unutar IRE130 došlo do ugradnje 1287 pb dugog elementa Mboumar iz porodice transpozona

mariner. Mboumar element se također preferencijalno ugrađuje uvijek na isto mjesto u sklopu IRE130, a može se naći i samostalno ugrađen usred niza satelitne DNA. U tom istraživanju nađen je i genomski fragment za koji autori pretpostavljaju da je, uzevši u obzir sva preferencijalna mjesta ugradnje pojedinih elemenata, nastao genetičkim rearanžmanom koji uključuje više vrsta ponovljenih sekvenci. Hipotetski početni fragmenet bi sadržavao niz satelitne DNA u sklopu kojeg se nalazi element IRE130, nešto dalje element Mboumar te kasnije u nastavku niza element Mboumar unutar elementa IRE130. Međusobnim sparivanjem 5' TIR sljeda prvog Mboumar elementa i 3' TIR sljeda drugog Mboumar elementa došlo bi do stvaranja omče i njihovog izrezivanja, što bi ostavilo samo jedan cjeloviti i jedan krnji element IRE130 u sklopu niza satelitne DNA. Autori čak nagađaju o mogućnosti da je visoka konzerviranost u sekvenci te satelitne DNA posljedica njezine mobilnosti putem ovih transponirajućih elemenata, popraćena poznatim homogenizacijskim mehanizmima.

Ono što je sigurno, ovo istraživanja pokazalo je blisku povezanost uzastopno i raspršeno ponovljenih sljedova, istovremeno postojanje jednog nukleotidnog sljeda u oba organizacijska tipa, kao i intenzivne interakcije među njima. Postojanje vrlo dinamičnih insercijskih, rekombinacijskih i homogenizacijskih procesa unutar svake skupine (kao i među njima), dovodi do stvaranja kompleksne „mreže“ uzastopno i raspršeno ponovljenih sekvenci, promovira širenje tih sljedova, uzrokuje značajne genetičke rearanžmane, čineći neospornim njihov utjecaj na oblikovanje i na veličinu, a time i na funkciju samog genoma.

5. Zaključci

1. Kod tri školjkaša koji su bili predmetom ovog istraživanja, *Donax trunculus*, *Ruditapes decussatus* i *Ruditapes philippinarum*, izoliran je niz uzastopno i raspršeno ponovljenih sekvenci. S obzirom na relativnu zastupljenost pojedinih tipova raspršeno ponovljenih sekvenci, kod sve 3 vrste uočena je blaga dominacija transpozona u odnosu na retrotranspozone.
2. U sklopu ovog istraživanja detaljno je opisan novi raspršeno ponovljeni element skupine MITE, DTC84, nađen u vrsti *D. trunculus*. Strukturne karakteristike koje opisuju ovaj element su: centralna regija koja sadrži kratke nizove uzastopnih ponavljanja s palindromskim sekvencama na njihovom kraju, mikrosatelitna regija utemeljena na tetranukletidnim ponavljanjima te konzervirani blokovi na krajevima elementa, u sklopu kojih su subterminalno i terminalno smješteni obrnuto ponovljeni motivi.
3. Element DTC84 izgrađuje 1% genoma ove vrste, a uzevši u obzir i odvedene varijante, procijenjena zastupljenost je i do 2%, što čini značajan doprinos ove raspršeno ponovljene sekvence genomu školjkaša.
4. Nezavisni evolucijski procesi upravljaju pojedinim strukturnim podjedinicama elementa DTC84, održavajući pritom visoko konzerviranim nukleotidne sljedove centralnih ponavljanja, palindromskih sekvenci i blokova na periferiji elementa kao i sam raspored centralnih ponavljanja, a istovremeno dopuštajući varijabilnost u broju centralnih i mikrosatelitnih ponavljanja.
5. Palindromske sekvence u sklopu centralnih ponavljanja vrlo vjerojatno predstavljaju motive odgovorne za insercijsko/delecijske rearanžmane ovih sljedova, dok sama centralna ponavljanja predstavljaju mogući izvor novih nizova satelitnih DNA.
6. Nove uzastopno ponovljene sekvence opisane u ovom radu uključuju dvije satelitne DNA. U vrsti *D. trunculus* otkrivena je prisutnost vrsno-specifične varijante satelitne DNA BIV160, široko rasprostranjene među školjkašima i pretpostavljene starosti oko 540 milijuna godina. Kod vrsta *R. decussatus* i *R. philippinarum* (porodica Veneridae) utvrđena je prisutnost satelitne DNA DTHS3, do sada nađene samo u vrsti *D. trunculus* (porodica Donacidae), pa je njezina minimalna starost procijenjena na 250 milijuna godina. Obje satelitne DNA pokazuju područja visoke konzerviranosti unutar nukleotidnog slijeda monomera, što ukazuje na potencijalne funkcionalne interakcije tih sljedova DNA.
7. Utvrđeno je da se pojedini nukleotidni sljedovi mogu istovremeno naći u 3 različita organizacijska oblika: kao centralna ponavljanja unutar mobilnih elemenata, kao

uzastopno ponovljeni monomeri satelitne DNA te kao zasebne kopije raspršene po genomu.

8. U sklopu ovog istraživanja po prvi put su unutar razreda školjkaša opisani mobilni elementi TRIM, uočeni kod vrsta *D. trunculus* i *R. philippinarum*.
9. Komparativne analize srodnih vrsta *R. decussatus* i *R. philippinarum* donose informaciju da ta dva školjkaša posjeduju niz zajedničkih uzastopno i raspršeno ponovljenih sljedova.
10. Nekoliko porodica satelitnih DNA vrste *D. trunculus* pokazuje sličnost u nukleotidnoj sekvenci s različitim tipovima mobilnih elemenata, ukazujući na blisku povezanost ta dva tipa ponovljenih sljedova.
11. Postojanje vrlo dinamičnih insercijskih, rekombinacijskih i homogenizacijskih procesa unutar svake skupine (kao i među njima) dovodi do stvaranja kompleksne „mreže“ uzastopno i raspršeno ponovljenih sekvenci, promovira širenje tih sljedova i uzrokuje genetičke rearanžmane, čineći neospornim njihov utjecaj na oblikovanje genoma.

6. Literatura

- [1] Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 2000; 408: 796–815.
- [2] Akasaki T, Nikaido M, Nishihara H, Tsuchiya K, Segawa S, Okada N. Characterization of a novel SINE superfamily from invertebrates: “Ceph-SINEs” from the genomes of squids and cuttlefish. *Gene* 2010; 454: 8–19.
- [3] Bachmann L, Rabb M, Schibel J, Sperlich D. Satellite DNA. U: Molecular techniques in taxonomy. (ured. Hewitt GM) Springer-Verlag Berlin, Heidelberg, 1991: 133–143.
- [4] Belancio V, Hedges D, Deininger P. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. *Genome research* 2008; 18: 343–58.
- [5] Benedum UM, Neitzel H, Sperling K, Bogenberger J, Fittier F. Organization and chromosomal distribution of a novel repetitive DNA component from *Muntiacus muntjak vaginalis* with a repeat length of more than 40 kb. *Chromosoma* 1986; 94: 267–272.
- [6] Bennetzen JL, Kellogg EA. Do Plants Have a One-Way Ticket to Genomic Obesity? *The Plant cell* 1997; 9: 1509–1514.
- [7] Biscotti MA, Canapa A, Olmo E i sur. Repetitive DNA, molecular cytogenetics and genome organization in the King scallop (*Pecten maximus*). *Gene* 2007; 406: 91–8.
- [8] Canapa A, Barucca M, Cerioni PN, Olmo E. A satellite DNA containing CENP-B box-like motifs is present in the antarctic scallop *Adamussium colbecki*. *Gene* 2000; 247: 175–80.
- [9] Capy P. Classification and nomenclature of retrotransposable elements. *Cytogenetic and genome research* 2005; 110: 457–61.
- [10] Cheng Z-J, Murata M. A centromeric tandem repeat family originating from a part of Ty3/gypsy-retroelement in wheat and its relatives. *Genetics* 2003; 164: 665–72.
- [11] Clabby C, Goswami U, Flavin F, Wilkins NP, Houghton J a, Powell R. Cloning, characterization and chromosomal location of a satellite DNA from the Pacific oyster, *Crassostrea gigas*. *Gene* 1996; 168: 205–9.
- [12] Coates BS, Kroemer JA, Sumerford D V., Hellmich RLL. A novel class of miniature inverted repeat transposable elements (MITEs) that contain hitchhiking (GTCY)(n) microsatellites. *Insect molecular biology* 2011; 20: 15–27.
- [13] Cohen S, Yacobi K, Segal D. Extrachromosomal circular DNA of tandemly repeated genomic sequences in *Drosophila*. *Genome research* 2003; 13: 1133–45.
- [14] Craig NL. Unity in Transposition Reactions. *Science* 1995; 270: 253–253.
- [15] Curcio MJ, Garfinkel DJ. New lines of host defense:inhibition of Ty1 retrotransposition by Fus3p and NER/TFIIH. *Trends in genetics* 1999; 15: 43–45.

- [16] Dover GA. Molecular drive in multigene families: How biological novelties arise, spread and are assimilated. *Trends in Genetics* 1986; 2: 159–165.
- [17] Eichler EE, Clark RA, She X. An assessment of the sequence gaps: unfinished business in a finished human genome. *Nature reviews. Genetics* 2004; 5: 345–54.
- [18] Feschotte C, Mouchès C. Evidence that a family of miniature inverted-repeat transposable elements (MITEs) from the *Arabidopsis thaliana* genome has arisen from a pogo-like DNA transposon. *Molecular biology and evolution* 2000; 17: 730–7.
- [19] Feschotte C, Pritham E. DNA Transposons and the Evolution of Eukaryotic Genomes. *Annu Rev Genet.* 2007; 41: 331–368.
- [20] Feschotte C, Zhang X, Wessler S. Miniature inverted-repeat transposable elements (MITEs) and their relationship with established DNA transposons. U: *Mobile DNA II.* (ured. Craig N) ASM Press, Washington, 2002: 1147–58.
- [21] Feschotte C. The contribution of transposable elements to the evolution of regulatory networks. *Nat Rev Genet* 2008; 9: 397–405.
- [22] Fry K, Salser W. Nucleotide sequences of HS- α satellite DNA from kangaroo rat *dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* 1977; 12: 1069–1084.
- [23] Gaffney PM, Pierce JC, Mackinley AG, Titchen DA, Glenn WK. Pearl, a novel family of putative transposable elements in bivalve mollusks. *Journal of molecular evolution* 2003; 56: 308–16.
- [24] Gladyshev EA, Arkhipova IR. Telomere-associated endonuclease-deficient Penelope-like retroelements in diverse eukaryotes. *PNAS* 2007; 104.
- [25] Gosling E. *Bivalve Molluscs: Biology, Ecology and Culture.* 2003.
- [26] Hall SE, Kettler G, Preuss D. Centromere Satellites From *Arabidopsis* Populations Maintenance of Conserved and Variable Domains. 2003;: 195–205.
- [27] Hamilton AT, Huntley S, Kim J, Branscomb E, Stubbs L. Lineage-specific Expansion of KRAB Zinc-finger Transcription Factor Genes: Implications for the Evolution of Vertebrate Regulatory Networks. *Cold Spring Harbor Symposia on Quantitative Biology* 2003; 68: 131–140.
- [28] Heikkinen E, Launonen V, Muller E, Bachmann L. The pvB370 BamHI satellite DNA family of the *Drosophila virilis* group and its evolutionary relation to mobile dispersed genetic pDv elements. *Journal of Molecular Evolution* 1995; 41.
- [29] Henikoff S, Ahmad K, Malik HS. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science (New York, N.Y.)* 2001; 293: 1098–102.

- [30] Hikosaka A, Kawahara A. Lineage-specific tandem repeats riding on a transposable element of MITE in *Xenopus* evolution: a new mechanism for creating simple sequence repeats. *Journal of molecular evolution* 2004; 59: 738–46.
- [31] Hurtado NS, Pérez-García C, Morán P, Pasantes JJ. Genetic and cytological evidence of hybridization between native *Ruditapes decussatus* and introduced *Ruditapes philippinarum* (Mollusca, Bivalvia, Veneridae) in NW Spain. *Aquaculture* 2011; 311: 123–128.
- [32] IRGSP. The map-based sequence of the rice genome. *Nature* 2005; 436: 793–800.
- [33] Izsvák Z, Ivics Z, Shimoda N, Mohn D, Okamoto H, Hackett PB. Short inverted-repeat transposable elements in teleost fish and implications for a mechanism of their amplification. *Journal of molecular evolution* 1999; 48: 13–21.
- [34] Jensen AC, Humphreys J, Caldow RWG, Grisley C, Dyrzynda PEJ. Naturalization of the Manila clam (*Tapes philippinarum*), an alien species, and establishment of a clam fishery within Poole Harbour, Dorset. *Journal of the Marine Biological Association of the UK* 2004; 84: 1069–1073.
- [35] Jiang J, Birchler JA, Parrott WA, Dawe, Kelly A. A molecular view of plant centromeres. *Trends in Plant Science* 2003; 8: 570–575.
- [36] Jiang N, Feschotte C, Zhang XY, Wessler SR. Miniature inverted-repeat transposable elements (MITEs) and their relationship with established DNA transposons. *Curr Opin Plant Biol* 2004; 7: 115–119.
- [37] John B. The biology of heterochromatin. U: *Heterochromatin: molecular and structural aspects*. (ured. Verma R) Cambridge University Press, Cambridge, 1988: 1–147.
- [38] Jordan IK, Rogozin IB, Glazko G V, Koonin E V. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends in genetics*: TIG 2003; 19: 68–72.
- [39] Jurka J. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proceedings of the National Academy of Sciences of the United States of America* 1997; 94: 1872–7.
- [40] Kalendar R, Antonius K, Smýkal P, Schulman AH. iPBS: a universal method for DNA fingerprinting and retrotransposon isolation. *TAG. Theoretical and applied genetics*. 2010; 121: 1419–30.
- [41] Kapitonov V V, Holmquist GP, Jurka J. L1 Repeat Is a Basic Unit of Heterochromatin Satellites in Cetaceans. *Molecular biology and evolution* 1998; 15: 611–612.
- [42] Kapitonov V V, Jurka J. Molecular paleontology of transposable elements from *Arabidopsis thaliana*. *Genetica* 1999; 107: 27–37.

- [43] Kapitonov V V, Jurka J. Molecular paleontology of transposable elements in the *Drosophila melanogaster* genome. *Proceedings of the National Academy of Sciences of the United States of America* 2003; 100: 6569–74.
- [44] Kapitonov VV, Jurka J. L1-56_XT family of frog non-LTR retrotransposons. *Rebase Reports* 2006; 6(12): 629–629.
- [45] Kidwell MG. Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 2002; 115: 49–63.
- [46] Kipling D, Warburton PE. Centromeres, CENP-B and Tigger too. *Trends in genetics* 1997; 13: 1–5.
- [47] Kleckner N. Regulation of transposition in bacteria. *Annual review of cell biology* 1990; 6: 297–327.
- [48] Kourtidis A, Drosopoulou E, Pantzartzi CN, Chintiroglou CC, Scouras ZG. Three new satellite sequences and a mobile element found inside HSP70 introns of the Mediterranean mussel (*Mytilus galloprovincialis*). *Genome* 2006; 49: 1451–8.
- [49] Kuhn GCS, Heslop-Harrison JS. Characterization and genomic organization of PERI, a repetitive DNA in the *Drosophila buzzatii* cluster related to DINE-1 transposable elements and highly abundant in the sex chromosomes. *Cytogenetic and genome research* 2011; 132: 79–88.
- [50] Kuhn GCS, Küttler H, Moreira-Filho O, Heslop-Harrison JS. The 1.688 repetitive DNA of *Drosophila*: concerted evolution at different genomic scales and association with genes. *Molecular biology and evolution* 2012; 29: 7–11.
- [51] van de Lagemaat LN, Landry J-R, Mager DL, Medstrand P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends in genetics* 2003; 19: 530–6.
- [52] Lander ES, Linton LM, Birren B i sur. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860–921.
- [53] Langdon T, Seago C, Mende M i sur. Retrotransposon evolution in diverse plant genomes. *Genetics* 2000; 156: 313–25.
- [54] Leitão A, Chaves R. Banding for Chromosomal Identification in Bivalves: A 20 -year History. *Dynamic Biochemistry, Process Biotechnology and Molecular Biology*, Global Science Books 2008; 2: 44–49.
- [55] Li J, Leung FC. A CR1 element is embedded in a novel tandem repeat (*Hinf*I repeat) within the chicken genome. *Genome* 2006; 49: 97–103.
- [56] Lister C, Jackson D, Martin C. Transposon-Induced Inversion in *Antirrhinum* Modifies *niwea* Gene Expression To Give a Novel Flower Color Pattern under the Control of *cycloidea radialis*. 1993; 5: 1541–1553.

- [57] López-Flores I, de la Herrán R, Garrido-Ramos M a, Boudry P, Ruiz-Rejón C, Ruiz-Rejón M. The molecular phylogeny of oysters based on a satellite DNA related to transposons. *Gene* 2004; 339: 181–8.
- [58] Maraia RJ, Sarrowa J. Alu-family SINE RNA: interacting proteins and pathways of expression. U: *The Impact of Short Interspersed Elements (SINEs) on the Host Genome.* (ured. Maraia R) Springer, New York 1995;: 163–196.
- [59] Marillonnet S, Wessler SR. Retrotransposon insertion into the maize waxy gene results in tissue-specific RNA processing. *The Plant cell* 1997; 9: 967–78.
- [60] Martinez-Lage A, Gonzalez-Tizon A, Mendez J. Characterization of different chromatin types in *Mytilus galloprovincialis* L. after C-banding, fluorochrome and restriction endonuclease treatments. *Heredity* 1994;: 242–249.
- [61] Martínez-Lage A, Rodríguez F, González-Tizón A, Prats E, Cornudella L, Méndez J. Comparative analysis of different satellite DNAs in four *Mytilus* species. *Genome* 2002; 45: 922–9.
- [62] Masumoto H, Masukata H, Muro Y, Nozaki N, Okazaki T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *The Journal of cell biology* 1989; 109: 1963–73.
- [63] McClintock B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci.* 1950; 36: 344–355.
- [64] Meštrović N, Mravinac B, Plohl M, Ugarković Đ, Bruvo B. Preliminary phylogeny of *Tribolium* beetles (Coleoptera: Tenebrionidae) resolved by combined analysis of mitochondrial genes. *Eur. J. Entomol.* 2006; 103: 709–715.
- [65] Meštrović N, Plohl M, Mravinac B, Ugarković D. Evolution of satellite DNAs from the genus *Palorus*--experimental evidence for the “library” hypothesis. *Molecular biology and evolution* 1998; 15: 1062–8.
- [66] Miller WJ, Nagel a, Bachmann J, Bachmann L. Evolutionary dynamics of the SGM transposon family in the *Drosophila obscura* species group. *Molecular biology and evolution* 2000; 17: 1597–609.
- [67] Moran J V, Arbor A. Chromosomes: Noncoding DNA (Including Satellite DNA). U: *ENCYCLOPEDIA OF LIFE SCIENCES.* John Wiley & Sons, 2005: 1–9.
- [68] Moran J V, DeBerardinis RJ, Kazazian HH. Exon shuffling by L1 retrotransposition. *Science* 1999; 283: 1530–4.
- [69] Mravinac B, Plohl M, Ugarković D. Conserved patterns in the evolution of *Tribolium* satellite DNAs. *Gene* 2004; 332: 169–77.
- [70] Muchmore ME, Moy GW, Swanson WJ, Vacquier VD. Direct sequencing of genomic DNA for characterization of a satellite DNA in five species of eastern Pacific abalone. *Molecular marine biology and biotechnology* 1998; 7: 1–6.

- [71] Nagaki K, Song J, Stupar RM i sur. Molecular and cytological analyses of large tracks of centromeric DNA reveal the structure and evolutionary dynamics of maize centromeres. *Genetics* 2003; 163: 759–70.
- [72] Nagy Z, Chandler M. Regulation of transposition in bacteria. *Research in microbiology* 2004; 155: 387–98.
- [73] Nakagawa H, Lee J-K, Hurwitz J i sur. Fission yeast CENP-B homologs nucleate centromeric heterochromatin by promoting heterochromatin-specific histone tail modifications. *Genes & development* 2002; 16: 1766–78.
- [74] Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu ZJ, Loftus B, Xi Z, Megy K, Grabherr M, Ren Q, Zdobnov EM, Lobo NF, Campbell KS, Brown SE, Bonaldo MF, Zhu J, Sinkins SP, Hogenkamp DG, Amedeo P, Arensburger P, Atkinson PW, Bidwell S, Biedler J, Birney E SD. Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* 2007; 316: 1718–1723.
- [75] Odierna G, Aprea G, Barucca M, Canapa a, Capriglione T, Olmo E. Karyology of the Antarctic scallop *Adamussium colbecki*, with some comments on the karyological evolution of pectinids. *Genetica* 2006; 127: 341–9.
- [76] Ohki R, Oishi M, Kiyama R. Preference of the recombination sites involved. *Nucleic Acids Research* 1995; 23: 4971–4977.
- [77] Ohzeki J, Nakano M, Okada T, Masumoto H. CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. *The Journal of cell biology* 2002; 159: 765–75.
- [78] Oosumi T, Garlick B, Belknap WR. Identification of putative nonautonomous transposable elements associated with several transposon families in *Caenorhabditis elegans*. *Journal of Molecular Evolution* 1996; 43: 11–18.
- [79] Palomeque T, Antonio Carrillo J, Muñoz-López M, Lorite P. Detection of a mariner-like element and a miniature inverted-repeat transposable element (MITE) associated with the heterochromatin from ants of the genus *Messor* and their possible involvement for satellite DNA evolution. *Gene* 2006; 371: 194–205.
- [80] Pardue ML, Debaryshe PG. *Drosophila* telomere transposons: genetically active elements in heterochromatin. *Genetica* 2000; 109: 45–52.
- [81] Passamonti M, Mantovani B, Scali V. Characterization of a highly repeated DNA family in tapetinae species (mollusca bivalvia: veneridae). *Zoological science* 1998; 15: 599–605.
- [82] Petrović V, Pérez-García C, Pasan tes JJ, Sato vić E, Prats E, Plo h IM. A GC-rich satellite DNA and karyology of the bivalve mollusk *Donax trunculus*: a dominance of GC-rich heterochromatin. *Cytogenetic and genome research* 2009; 124: 63–71.
- [83] Petrović V, Plohl M. Sequence divergence and conservation in organizationally distinct subfamilies of *Donax trunculus* satellite DNA. *Gene* 2005; 362: 37–43.

- [84] Plasterk RHA. Resident aliens. *Trends in Genetics* 1999; 15.
- [85] Plohl M, Bruvo B, Meštrović N, Mravinac B, Petrović V, Durajlija-Žinić, Sonja Ugarković Đ. Satellite DNA sequences in centromeric heterochromatin. *Periodicum Biologorum* 2004; 106: 95–102.
- [86] Plohl M, Cornudella L. Characterization of a complex satellite DNA in the mollusc *Donax trunculus*: analysis of sequence variations and divergence. *Gene* 1996; 169: 157–64.
- [87] Plohl M, Cornudella L. Characterization of interrelated sequence motifs in four satellite DNAs and their distribution in the genome of the mollusc *Donax trunculus*. *Journal of molecular evolution* 1997; 44: 189–98.
- [88] Plohl M, Luchetti A, Mestrovic N, Mantovani B. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* 2008; 409: 72–82.
- [89] Plohl M, Meštrović N, Mravinac B. Satellite DNA evolution. U: *Genome dynamics*. (ured. Garrido- Ramos M) Karger AG, Basel, 2012: 126–52.
- [90] Plohl M, Petrović V, Luchetti a i sur. Long-term conservation vs high sequence divergence: the case of an extraordinarily old satellite DNA in bivalve mollusks. *Heredity* 2010; 104: 543–51.
- [91] Plohl M, Prats E, Martinez-Lage A, Gonzalez-Tizon A, Mendez J, Cornudella L. Telomeric localization of the vertebrate-type hexamer repeat, (TTAGGG)_n, in the wedgeshell clam *Donax trunculus* and other marine invertebrate genomes. *The Journal of biological chemistry* 2002; 277: 19839–46.
- [92] Pritham EJ, Feschotte C. Massive amplification of rolling-circle transposons in the lineage of the bat *Myotis lucifugus*. *Proceedings of the National Academy of Sciences of the United States of America* 2007; 104: 1895–900.
- [93] Pritham EJ, Putliwala T, Feschotte C. Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene* 2007; 390: 3–17.
- [94] Ramsay L, Macaulay M, Cardle L i sur. Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. *The Plant journal: for cell and molecular biology* 1999; 17: 415–25.
- [95] Robertson HM. Members of the pogo superfamily of DNA-mediated transposons in the human genome. *MGG Molecular & General Genetics* 1996; 252: 761–766.
- [96] Ruiz-Lara S, Prats E, Sainz J, Cornudella L. Cloning and characterization of a highly conserved satellite DNA from the mollusc *Mytilus edulis*. *Gene* 1992; 117: 237–42.
- [97] Saavedra C, Bachère E. Bivalve genomics. *Aquaculture* 2006; 256: 1–14.

- [98] SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL. The paleontology of intergene retrotransposons of maize. *Nature Genetics* 1998; 20: 43–45.
- [99] Schramke V, Allshire R. Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. *Science (New York, N.Y.)* 2003; 301: 1069–74.
- [100] Schueler MG, Higgins a W, Rudd MK, Gustashaw K, Willard HF. Genomic and genetic definition of a functional human centromere. *Science (New York, N.Y.)* 2001; 294: 109–15.
- [101] Sigwart JD, Sutton MD. Deep molluscan phylogeny: synthesis of palaeontological and neontological data. *Proceedings. Biological sciences / The Royal Society* 2007; 274: 2413–9.
- [102] Smith G. Evolution of repeated DNA sequences by unequal crossover. *Science* 1976; 191: 528–535.
- [103] Sullivan KF, Glass CA. CENP-B is a highly conserved mammalian centromere protein with homology to the helix-loop-helix family of proteins. *Chromosoma* 1991; 100: 360–370.
- [104] Sun X, Le HD, Wahlstrom JM, Karpen GH. Sequence analysis of a functional *Drosophila* centromere. *Genome research* 2003; 13: 182–94.
- [105] Swamy L, Wessler SR. Inverted Repeat Transposable Elements (MITEs). *Genetics* 2003; 758: 747–758.
- [106] Takeuchi T, Kawashima T, Koyanagi R i sur. Draft genome of the pearl oyster *Pinctada fucata*: a platform for understanding bivalve biology. *DNA research an international journal for rapid publication of reports on genes and genomes* 2012; 19: 117–30.
- [107] Tang M, Waterman M, Yooseph S. Zinc finger gene clusters and tandem gene duplication. *Journal of computational molecular cell biology* 2002; 9: 429–46.
- [108] Tanguy A, Bierne N, Saavedra C i sur. Increasing genomic information in bivalves through new EST collections in four species: development of new genetic markers for environmental studies and genome evolution. *Gene* 2008; 408: 27–36.
- [109] Tek AL, Song J, Macas J, Jiang J. Sobo, a recently amplified satellite repeat of potato, and its implications for the origin of tandemly repeated sequences. *Genetics* 2005; 170: 1231–8.
- [110] Tudor M, Lobočka M, Goodell M, Pettitt J, O'Hare K. The pogo transposable element family of *Drosophila melanogaster*. *MGG Molecular & General Genetics* 1992; 232: 126–134.
- [111] Ugarkovic D. Functional elements residing within satellite DNAs. *EMBO reports* 2005; 6: 1035–9.

- [112] Ugarković Đ, Plohl M. Variation in satellite DNA profiles-causes and effects. *The EMBO Journal* 2002; 21: 5955–5959.
- [113] Végh Z, Vincze E, Kadirov R, Tóth G, Kiss GB. The nucleotide sequence of a nodule-specific gene, Nms-25 of *Medicago sativa*: its primary evolution via exon-shuffling and retrotransposon-mediated DNA rearrangements. *Plant molecular biology* 1990; 15: 295–306.
- [114] Verdel A, Jia S, Gerber S i sur. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science (New York, N.Y.)* 2004; 303: 672–6.
- [115] Volpe TA, Kidner C, Hall IM, Teng G, Grewal SIS, Martienssen R a. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science (New York, N.Y.)* 2002; 297: 1833–7.
- [116] Wang Y, Xu Z, Guo X. A Centromeric Satellite Sequence in the Pacific Oyster (*Crassostrea gigas* Thunberg) Identified by Fluorescence In Situ Hybridization. *Marine Biotechnology* 2001; 3: 486–492.
- [117] Warburton PE, Waye JS, Willard HF. Nonrandom localization of recombination events in human alpha satellite repeat unit variants: implications for higher-order structural characteristics within centromeric heterochromatin. *Molecular and cellular biology* 1993; 13: 6520–9.
- [118] Wilder J, Hollocher H. Mobile elements and the genesis of microsatellites in dipterans. *Molecular biology and evolution* 2001; 18: 384–92.
- [119] Witte CP, Le QH, Bureau T, Kumar a. Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. *Proceedings of the National Academy of Sciences of the United States of America* 2001; 98: 13778–83.
- [120] Yang H-P, Barbash D a. Abundant and species-specific DINE-1 transposable elements in 12 *Drosophila* genomes. *Genome biology* 2008; 9: R39.
- [121] Yang H-P, Barbash DA. Abundant and species-specific DINE-1 transposable elements in 12 *Drosophila* genomes. *Genome biology* 2008; 9: R39.
- [122] Yang T-J, Kwon S-J, Choi B-S i sur. Characterization of terminal-repeat retrotransposon in miniature (TRIM) in Brassica relatives. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* 2007; 114: 627–36.
- [123] Yang Z, Boffelli D, Boonmark N, Schwartz K, Lawn R. Apolipoprotein(a) gene enhancer resides within a LINE element. *The Journal of biological chemistry* 1998; 273: 891–7.
- [124] Yoshida M, Ishikura Y, Moritaki T i sur. Genome structure analysis of molluscs revealed whole genome duplication and lineage specific repeat variation. *Gene* 2011; 483: 63–71.

- [125] Zhang G, Fang X, Guo X i sur. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* 2012; 490: 49–54.
- [126] Zhang X, Feschotte C, Zhang Q, Jiang N, Eggleston WB, Wessler SR. P instability factor: an active maize transposon system associated with the amplification of Tourist-like MITEs and a new superfamily of transposases. *Proceedings of the National Academy of Sciences of the United States of America* 2001; 98: 12572–7.
- [127] Zhang Y, Huang Y, Zhang L i sur. Structural features of the rice chromosome 4 centromere. *Nucleic acids research* 2004; 32: 2023–30.
- [128] Zhou Y, Cahan SH. A novel family of terminal-repeat retrotransposon in miniature (TRIM) in the genome of the red harvester ant, *Pogonomyrmex barbatus*. *PloS one* 2012; 7: e53401.

7. Sažetak

Značajan dio eukariotskog genetičkog materijala čine uzastopno i raspršeno ponovljene sekvence DNA. Među njima, satelitne DNA predstavljaju dominantnu skupinu uzastopno ponovljenih sekvenci koje svojim rasporedom definiraju arhitekturu kromosoma, dok transponirajući elementi svojim širenjem imaju ključnu ulogu u genomskoj evoluciji. Školjkaši predstavljaju veliku skupinu beskralješnjaka izuzetnog ekološkog značaja i gospodarskog potencijala, međutim istraživanja ovih vrsta na razini gena i genoma uglavnom su zapostavljena. Kako bi se omogućio potpuniji uvid u strukturu i organizaciju genoma vrsta ove skupine s obzirom na ponovljene sekvence DNA te kako bi upotunili naše spoznaje o samoj prirodi tih sljedova, u ovom sam radu istražila 3 vrste školjkaša od ekološke i komercijalne važnosti: *Donax trunculus*, *Ruditapes decussatus* i *Ruditapes philippinarum*. U tu svrhu konstruirane su genomske biblioteke dotičnih vrsta koje su u potrazi za različitim tipovima ponovljenih sekvenci pretražene hibridizacijskim metodama, uz korištenje ukupnih genomskih DNA kao hibridizacijskih sonda. Kod sve tri vrste školjkaša izdvojen je niz različitih uzastopno i raspršeno ponovljenih sljedova, a gledajući relativnu zastupljenost raspršeno ponovljenih elemenata uočena je blaga dominacija transpozona u odnosu na retrotranspozone. U vrsti *Donax trunculus* opisan je minijaturni transponirajući element s obrnutim ponavljanjima (MITE), nazvan DTC84, koji dijeli modularnu organizaciju, a donekle i nukleotidnu sličnost s elementima *pearl* iz kamenice *Crassostrea virginica*. Ovi elementi sadrže kratki niz centralnih ponavljanja, a kod DTC84 je po prvi put opažen uređeni raspored njihovih varijanti. Palindromske sekvence prisutne unutar ~160 pb dugih jedinica centralnih ponavljanja predstavljaju motive koji bi mogli biti odgovorni za njihove insercijske i/ili delecijске rearanžmane, a sami nizovi centralnih ponavljanja mogu predstavljati ishodišta satelitnih DNA. U prilog ovakvoj pretpostavci govori činjenica da nekoliko porodica satelitnih DNA vrste *D. trunculus* pokazuje sličnost s različitim mobilnim elementima, ukazujući na blisku povezanost ta dva tipa ponovljenih sekvenci. Tako je unutar vrste *Donax trunculus* otkrivena vrsno-specifična varijanta satelitne DNA BIV160, široko rasprostranjene među školjkašima i pretpostavljene starosti oko 540 milijuna godina, a koja pokazuje nukleotidnu sličnost s centralnim ponavljanjima *pearl* elementa CvA iz kamenice *C. virginica*. Analiza ponovljenih sekvenci DNA ovih vrsta nadalje je pokazala da se određeni nukleotidni slijed može istovremeno naći u 3 različita organizacijska oblika: kao centralna ponavljanja unutar mobilnih elemenata, kao uzastopno ponovljeni monomeri satelitne DNA te u obliku zasebnih kopija, raspršenih u genomu. U genomima vrsta *R. decussatus* i *R. philippinarum* (porodica Veneridae) utvrđena je prisutnost satelitne DNA DTHS3, do sada nađene samo u vrsti *D. trunculus* (porodica Donacidae), pa je pretpostavljena minimalna

starost ove sekvence 250 milijuna godina. Također, po prvi put je unutar razreda školjkaša uočen i opisan mobilni element TRIM, prisutan kod vrsta *D. trunculus* i *R. philippinarum*. Usporedbom izdvojenih ponovljenih sekvenci blisko srodnih vrsta *R. decussatus* i *R. philippinarum* uočeno je da one dijele niz zajedničkih uzastopno i raspršeno ponovljenih sljedova. Srodnost satelitnih DNA i mobilnih elemenata ovih vrsta ukazuje da vrlo dinamični insercijski, rekombinacijski i homogenizacijski procesi povezuju ove sekvence u jedinstvenu kompleksnu mrežu koja oblikuje sveukupni sadržaj ponovljenih sekvenci u genomu.

8. Summary

Significant part of most eukaryotic genomes constitute tandem and interspersed repetitive DNA sequences. Among them, satellite DNAs form a dominant group that defines chromosomal architecture, while expansion of transposable elements plays the key role in genome evolution. Bivalves are a large group of invertebrate organisms with big ecological importance and economical potential but studies of these species at the level of genes and genomes are still underrepresented. In order to provide a better insight into the structure and organization of bivalve genomes, as well as to amend our knowledge about the role and evolution of repetitive DNA sequences, I explored 3 bivalve species of ecological and commercial importance: *Donax trunculus*, *Ruditapes decussatus* and *Ruditapes philippinarum*. For this purpose genomic libraries were constructed and hybridization-based screening was performed, using total genomic DNA as hybridization probes. A number of different tandem and interspersed repetitive DNA sequences were isolated. Looking at the relative abundance of interspersed genetic elements, all 3 bivalves exhibit mild dominance of transposons in respect to retrotransposons. New miniature inverted-repeat transposable element, DTC84, was characterized in the clam *Donax trunculus*. It shares modular organization, and, to some extent, nucleotide sequence similarity with *pearl* elements from the oyster *Crassostrea virginica*. These elements have a central region built of tandem repeats, while a unique feature, described for the first time in this work, is fixed position of variants inside the central repeat array of DTC84. Palindromes residing within ~160 bp long central repeat units represent motifs that might be responsible for insertion and/or deletion events of these sequences. In that way, expansion of central repeats may give rise to satellite DNA arrays. In support of such assumption speaks the fact that several satellite DNA families from *D. trunculus* show similarity to various mobile elements, indicating a close relationship between these two types of repetitive sequences. Thus, the species-specific variant of BIV160 satellite DNA, broadly distributed among bivalve mollusks and presumably 540 million years old, was detected in *D. trunculus* genome, showing sequence similarity to central repeats of the oyster *C. virginica* *pearl* element CvA. Further analysis revealed possibility for the same nucleotide sequence to be present in three different organizational forms, as core repeats of mobile elements, as arrays of satellite DNA monomers, and as individual copies dispersed throughout the genome. DTHS3 satellite DNA, detected until now only in *D. trunculus* (family Donacidae), during this work was observed in *R. decussatus* and *R. philippinarum* genomes (family Veneridae) as well, what makes this sequence presumably 250 million years old. This work also brings a first report of TRIM mobile element in bivalve species, detected within *D. trunculus* and *R. philippinarum* genomes. Comparison of repetitive DNA sequences

originating from related species *R. decussatus* and *R. philippinarum* disclosed a set of tandem and interspersed repeats shared among them. Close relationships between satellite DNAs and mobile elements in explored bivalve species indicate that highly dynamic insertion, homogenization and recombination processes interlink these sequences into a unique complex network that shapes genomic repetitive environment.

9. Popis kratica

cDNA	<i>complementary DNA</i>
CDP-Star	dinatrij 2-kloro-5-(4-metoksi-spiro{1,2-dioksetan-3,2'-(5'-kloro) triciklo [3.3.1.1 ^{3,7}]dekan}-4-il)-1-fenil fosfat
DIG	<i>Digoxigenin</i>
DNA	deoksiribonukleinska kiselina
EDTA	Etilen diamin tetraoctena kiselina
EST	<i>Expressed sequence tag</i>
FISH	<i>Fluorescent in Situ Hybridization</i>
Gb	giga baza
HOR	<i>Higher-order repeat</i>
LINE	<i>Long Interspersed Nuclear Element</i>
LTR	<i>Long terminal repeats</i>
MITE	<i>Miniature inverted-repeat transposable element</i>
MIR	<i>Mammalian interspersed elements</i>
ORF	<i>Open reading frame</i>
pb	parova baza
PBS	<i>Primer binding site</i>
PCR	<i>Polymerase chain reaction</i>
PPT	<i>Polypurine tract</i>
RNA	ribonukleinska kiselina
SDS	<i>Sodium dodecyl sulfate</i>
SINE	<i>Short Interspersed Nuclear Element</i>
SRP	<i>Signal recognition particle</i>
TDR	<i>Terminal direct repeat</i>
TIR	<i>Terminal inverted repeat</i>
TRIM	<i>Terminal-repeat retrotransposons in miniature</i>
TSD	<i>Target site duplication</i>

UTR	<i>Untranslated region</i>
TAE	<i>Tris acetic acid EDTA</i>
TE	<i>Tris-EDTA</i>
UV	<i>ultraviolet</i>

10. Životopis

Rođena sam 1. lipnja 1982. godine u Zagrebu, gdje sam pohađala osnovnu školu klasičnog usmjerenja i II opću gimnaziju. 2006. godine završila sam studij biologije na Prirodoslovno-matematičkom fakultetu Sveučilišta u Zagrebu, smjer: molekularna biologija. Diplomski rad sam izradila u Laboratoriju za molekularnu i staničnu biologiju Instituta Ruđer Bošković u Zagrebu, pod vodstvom dr.sc. Ivica Rubelja. Od lipnja 2007. godine zaposlena sam na Institutu Ruđer Bošković u Laboratoriju za strukturu i funkciju heterokromatina u statusu znanstvenog novaka pod vodstvom dr.sc. Miroslava Plohla. Sveučilišni poslijediplomski interdisciplinarni doktorski studij Molekularne bioznanosti Sveučilišta Josipa Jurja Strossmayera u Osijeku, Sveučilišta u Dubrovniku i Instituta Ruđer Bošković upisala sam 2008. godine. U sklopu svojeg znanstveno-istraživačkog rada gostovala sam u laboratoriju dr.sc. Barbare Mantovani (Laboratorij za molekularnu zoologiju Sveučilišta u Bologni, Italija) u trajanju 2 tjedna, te u laboratoriju dr.sc. Juana Pasantesa (Laboratorij za populacijsku genetiku i citogenetiku Sveučilišta u Vigu, Španjolska) u trajanju 3 tjedna. Članica sam SEB (Society for Experimental Biology) udruženja. Prisustvovala sam na nekoliko domaćih i stranih znanstvenih skupova i radionica:

- Heterochromatin Structure And Function: From Repetitive DNA Sequences To Epigenetics, ESF Exploratory Workshop, Donja Stubica, Hrvatska (2008)
- Konferencija “50 godina molekularne biologije u Hrvatskoj”, Zagreb, Hrvatska (2008)
- Simpozij „EMBO Young Scientists Forum“, Zagreb, Hrvatska (2009)
- Tečaj „Introduction to Bioinformatics“, Zagreb, Hrvatska (2009)
- 10. hrvatski biološki kongres, Osijek, Hrvatska (2009)
- PhD Student Symposium “The Architecture of Life”, Barcelona, Španjolska (2009)
- 6th Congress of the European Malacological Societies, Vitoria-Gasteiz, Španjolska (2011)
- 18th International Chromosome Conference, Manchester, UK (2011)
- Metodološki tečajevi u biologiji i medicini „Molekularna filogenija“, Zagreb, Hrvatska (2011)

Popis publikacija:

Plohl, Miroslav; Petrović, Vlatka; Luchetti, Andrea; Ricci, Andrea; Šatović, Eva; Passamonti, Marco; Mantovani, Barbara. **Long-term conservation vs. high sequence divergence : the case of an extraordinarily old satellite DNA in bivalve mollusks.** // *Heredity*. 104 (2010), 6; 543-551

Petrović, Vlatka; Perez-Garcia, Concepcion; Pasantes, Juan; Šatović, Eva; Prats, Eva; Plohl, Miroslav. **A GC-rich satellite DNA and karyology of the bivalve mollusc *Donax trunculus*: a dominance of GC-rich heterochromatin.** // *Cytogenetics and Genome Research*. 124 (2009), 1; 63-71