

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO - MATEMATIČKI FAKULTET
BIOLOŠKI ODSJEK

OTKRIĆE NOVOG TIPA
POKRETNIH GENETIČKIH ELEMENATA

DISCOVERY OF A NEW TYPE
OF MOBILE GENETIC ELEMENTS

Seminarski rad

Lucija Papa
Preddiplomski studij molekularne biologije
Undergraduate studies of Molecular Biology
Mentorica: doc. dr. sc. Ivana Ivančić Baće

Zagreb, 2014.

SADRŽAJ

1. Uvod	1
2. Otkriće kaspozona	3
3. Struktura kaspozona	4
3.1 Geni za nukleaze	4
3.2 Geni za proteine uključene u replikaciju DNA	4
3.3 Geni za male DNA-vezujuće proteine	5
3.4 Dodatni geni	5
3.5 Krajevi kaspozona	5
4. Mehanizam izrezivanja i insercije kaspozona	7
5. Skupine kaspozona	10
5.1 Porodica 1	10
5.2 Porodica 2	10
5.3 Porodica 3	12
6. Sličnost kaspozona i drugih pokretnih genetičkih elemenata	13
7. Evolucijski značaj proteina Cas1	17
8. Zaključak	18
9. Literatura	19
10. Sažetak	20
11. Summary	21

1. UVOD

Pokretne genetičke elemente otkrila je četrdesetih godina prošlog stoljeća Barbara McClintock radeći klasične genetičke pokuse na kukuruzu. Za ovo je otkriće 1983. godine nagrađena Nobelovom nagradom u području medicine (Lodish i sur. 2000).

Pokretni genetički elementi su definirane DNA sekvence različitih duljina (od 1 do nekoliko 1000 kb) koje se ugrađuju u kromosome domaćinskih stanica, te imaju sposobnost promjene svoje genetske lokacije: mogu se pokretati unutar genoma u kojem se nalaze (intracelularna pokretljivost) ili prenijeti u genom druge stanice (ekstracelularna pokretljivost). Na ovaj način, s obzirom na mjesto u genomu u koje se ugrađuju, mogu izravno utjecati na aktivnost gena. Pokretni genetički elementi često sadrže sekvence koje kodiraju proteine potrebne za vlastiti prijenos, ugradnju u genom domaćina i rekombinacije unutar genoma (Leplae i sur. 2004).

U pokretne genetičke elemente spadaju DNA transpozoni, retrotranspozoni, insercijske sekvence, virusne sekvence i plazmidi (Leplae i sur. 2004, Lodish i sur. 2000), te samoizrezujući introni skupina I i II (Martinez-Abarca i Toro 2002). Sekvenciranjem i usporedbom sve većeg broja genoma različitih organizama otkriva se također sve više novih pokretnih genetičkih elemenata, njihovih skupina, kao i "kimernih genetičkih elemenata". Ti "kimerni elementi" sadrže značajke više različitih skupina, te dijele zajedničke gene ili grupe gena (Leplae i sur. 2004).

Pokretni genetički elementi sadržani su u stanicama i prokariota i eukariota, te čine velik udio domaćinskog genoma. Smatra se da u bakterijskim stanicama pokretni genetički elementi i njihovi ostaci čine oko 30% genoma, u čovjeka 35% do 50% (Krupovic i sur. 2014), dok kod kukuruza njihov udio doseže i preko 85% ukupnog genoma (Tenailon i sur. 2011). Ugradnja pokretnih genetičkih elemenata u genom domaćina može uzrokovati štetu (na primjer, ugradnja pokretnog genetičkog elementa usred nekog gena dovodi do njegove inaktivacije), korist (neki pokretni genetički elementi mogu nositi gene za proizvodnju toksina ili otpornost na antibiotike) ili ne uzrokovati nikakvu promjenu (Krupovic i sur. 2014). Smatra se da su zbog svoje uloge u horizontalnom prijenosu gena i modifikacijama domaćinskog genoma pokretni genetički elementi jedni od glavnih pokretača evolucije (Leplae i sur. 2004).

Kaspozoni su novootkriveni tip samo-sintetizirajućih pokretnih genetičkih elemenata, pronađenih kod prokariotskih organizama (bakterija i arheja). Za sada su pronađeni i potvrđeni analizom genskih i proteinskih sekvenci kod ukupno 19 vrsta prokariota, a dijele se u porodice 1, 2 i 3. Trenutno su kaspozoni jedini poznati pokretni genetički elementi koji za integraciju u genom domaćina koriste endonukleazu Cas1, prema čemu su dobili ime. Svi poznati kaspozoni, osim gena za endonukleazu Cas1, obavezno sadrže i gen za vlastitu DNA polimerazu PolB, najvjerojatnije uključenu u njihovu replikaciju. Zbog svojstva da sami kodiraju enzim potreban za vlastito umnažanje, spadaju u skupinu samo-sintetizirajućih pokretnih genetičkih elemenata. Sastav drugih gena u kaspozonima znatno varira. Omeđeni su terminalnim repeticijama, te direktnim repeticijama koje nastaju kao posljedica njihova ugrađivanja u genom domaćina.

Kao kriterij za prepoznavanje određene genske sekvence kao kaspozona nužno je postojanje gena za endonukleazu Cas1 i DNA polimerazu PolB u neposrednoj blizini (Krupovic i sur. 2014).

2. OTKRIĆE KASPOZONA

Kaspozoni su otkriveni tijekom genomske analize gena *cas*. Geni *cas* su do tada bili povezivani prvenstveno sa sustavom CRISPR-Cas, koji omogućuje postojanje specifične imunosti kod prokariota. Ključan enzim u sustavu CRISPR-Cas je endonukleaza kodirana genom *casI*. Međutim, analiza je ukazala na postojanje još dvije filogenetski zasebne skupine gena *casI* koje nisu povezane sa sustavom CRISPR-Cas, te su nazvane Cas1-solo.

Geni *casI* koji pripadaju prvoj skupini Cas1-solo pronađeni su samo u arhejama reda *Methanomicrobiales* i ne pokazuju dokaze o postojanju horizontalnog prijenosa gena. Također ne kodiraju sve potrebne komponente katalitičkog mjesta koje su kod proteina Cas1 sustava CRISPR-Cas nužne za endonukleaznu aktivnost, te se smatra kako produkti ovih gena ne mogu biti aktivne endonukleaze.

S druge strane, geni *casI* koji pripadaju drugoj skupini Cas1-solo pronađeni su kod većeg broja različitih, međusobno nepovezanih vrsta bakterija i arheja, što upućuje na postojanje horizontalnog prijenosa gena. Najčešće se nalaze u očuvanom području, zajedno s genom za DNA polimerazu PolB (sličnu polimerazama PolB nekih pokretnih genetičkih elemenata), te genima za endonukleazu HNH i dva manja proteina sa uzvojnica-okret-uzvojnica domenama. Proteini Cas1 dobiveni translacijom gena *casI* druge skupine Cas1-solo imaju očuvano katalitičko mjesto, te se radi o aktivnim endonukleazama.

Poput DNA transpozona, ova očuvana genska područja su u najvećem broju slučajeva omeđena invertnim i direktnim repeticijama, te su veličinom slična eukariotskim DNA transpozonomima iz skupine Polintona/Mavericka. Zbog velikog broja sličnosti između pronađenih očuvanih sekvenci i različitih pokretnih elemenata, zaključeno je da se radi o novom tipu pokretnih genetičkih elemenata, koji su zbog prisustva gena za endonukleazu Cas1 nazvani kaspozonomima.

Kaspozoni su nađeni kod nekih arheja reda *Thermoplasmatales* i koljena Thaumarchaeota, ali najviše kod arheja razreda *Methanomicrobia*; te kod bakterijskih vrsti *Streptomyces albulus* CCRC 11814, *Henriciella marina* DSM 19595, *Nitrosomonas* sp. AL212 i u genomskim ostacima neklasificirane bakterije *Candidatus Acetothermum autotrophicum* (Krupovic i sur. 2014).

3. STRUKTURA KASPOZONA

Iako sastav njihovih gena znatno varira, te postoje različite grane kaspozona, svi za sada ispitani kaspozoni sadrže gene za proteine koji ostvaruju interakcije s DNA, raspoređene u tri skupine. U prvu skupinu gena spadaju geni za endonukleaze Cas1 i dodatne nukleaze koje mogu i ne moraju biti prisutne; u drugu skupinu spadaju geni za DNA polimerazu PolB i eventualne dodatne proteine uključene u replikaciju DNA; a u treću geni za različite male DNA-vezujuće proteine. Neki kaspozoni sadrže i dodatne gene koji ne ostvaruju nužno interakcije s DNA.

3.1 GENI ZA NUKLEAZE

Osim gena za endonukleazu Cas1, kojeg sadrže svi kaspozoni, neki kaspozoni nose i gene za dodatne nukleaze koje bi potencijalno mogle sudjelovati u njihovom izrezivanju i integraciji. Ovi dodatni geni mogu uključivati gen za nukleazu HNH (koja vjerojatno sa Cas1 sudjeluje u izrezivanju i inserciji); gene za apurinske/apirimidinske endonukleaze (pronađene kod MetLum-C1 i MetPsy-C1); GIY-YIG nukleazu; i uracil-DNA-glikozilazu koja bi mogla sudjelovati u popravku krajeva DNA nakon insercije kaspozona.

Katalitičko središte Cas1 endonukleaze sastoji se od dva aspartata na pozicijama 218 i 221, histidina na poziciji 208 i glutamata na poziciji 141. Aktivnost proteina Cas1 dokazana je na DNA supstratima različitih konformacija.

3.2 GENI ZA PROTEINE UKLJUČENE U REPLIKACIJU DNA

Svi dosad poznati kaspozoni obavezno nose i gen za jedan od dva tipa DNA polimeraze PolB. Mnogi od njih sadrže i sekvence za različite proteine koji bi mogli pomagati prilikom replikacije DNA: na primjer, AceAut-C1 kaspozon kodira i helikazu sličnu HerA, protein koji bi mogao biti postavljač helikaze sličan DnaC, te dodatni protein s AAA i ATPaznom domenom koja je nađena kod mnogih helikaza. Kod HenMar-C1 nađen je jedinstveni fuzijski protein, kojem je N-terminusna domena srodna proteinima sličnima Cas1 i ima nukleaznu aktivnost, dok C-terminusna domena ima aktivnost helikaze.

3.3 GENI ZA MALE DNA-VEZUJUĆE PROTEINE

Većina kaspozona nosi gene za različite kombinacije malih DNA-vezujućih proteina. Ovi proteini mogu sadržavati domene s motivima uzvojnica-okret-uzvojnica (helix-turn-helix), vrpca-uzvojnica-uzvojnica (ribbon-helix-helix); ili cink-vezujuće domene. Vjerojatno pomažu nukleazama u izrezivanju i integraciji kaspozona, prilikom replikacije kaspozonske DNA, ili u regulaciji kaspozonskih gena. Neki mali DNA-vezujući proteini kodirani kaspozonomima mogli bi imati ulogu u regulaciji ekspresije gena domaćina.

3.4 DODATNI GENI

Za razliku od gena koji pripadaju u neku od osnovnih skupina koje sadrže svi kaspozoni, dodatni geni, osim interakcija s DNA, mogu ostvarivati i interakcije s RNA, te biti uključeni u metaboličke procese ili mehanizme obrane stanice.

U tri vrste kaspozona nađen je protein sličan Sir-2. Sir-2 je enzim koji kod arheja deacetilira kromosomski protein Alba, te se smatra da bi kaspozonski protein sličan Sir-2 mogao imati sličnu funkciju. Kod nekih vrsta kaspozona pronađeni su geni koji bi mogli kodirati za DNA metiltransferaze, restrikcijske endonukleaze, serinske i treoninske kinaze, RNA-vezujući protein sličan Sm i druge. U kaspozonu AmiAut-C1 nađen je fuzijski protein koji u N-terminalnoj domeni ima acetiltransferaznu aktivnost, a u C-terminalnoj domeni kveozin tRNA-riboziltransferaznu aktivnost.

3.5 KRAJEVI KASPOZONA

Kaspozoni porodice 1 nose međusobno raznolike gene za terminalne proteine koji bi mogli poslužiti kao početnice prilikom replikacije DNA. Sekvence koje kodiraju terminalne proteine nalaze se na očuvanim pozicijama u blizini i uzvodno od gena *polB*, ili terminalni proteini postoje kao privjesak na N-terminusu PolB.

Kaspozonski geni su u pravilu uvijek omeđeni invertnim i direktnim repeticijama. Duljine invertnih repeticija variraju od 25 do 602 nukleotida (najčešće 56 nukleotida), te ponekad mogu sadržavati unutrašnje palindromske sekvence. Invertne repeticije nisu pronađene kod tri kaspozona nađena u genomu vrste *Methanococcoides burtonii* DSM 6242 (MetBur-C1, MetBur-C2 i MetBur-C3).

Direktne repeticije nastaju kao posljedica integracije kaspozona u ciljnu sekvencu i variraju u duljini od 1 do 27 nukleotida. Kod određenog dijela kaspozona uočeno je djelomično preklapanje jedne ili obje invertne repeticije s direktnim. Pretpostavlja se da preklapanja nastaju zbog toga što nakon izrezivanja ovi kaspozoni vjerojatno sadrže terminalne jednolančane privjeske, koji su djelomično komplementarni sekvencama na krajevima zarezanog ciljnog mjesta. Kaspozoni kod kojih ovakva preklapanja nisu nađena vjerojatno nakon izrezivanja imaju tupe krajeve.

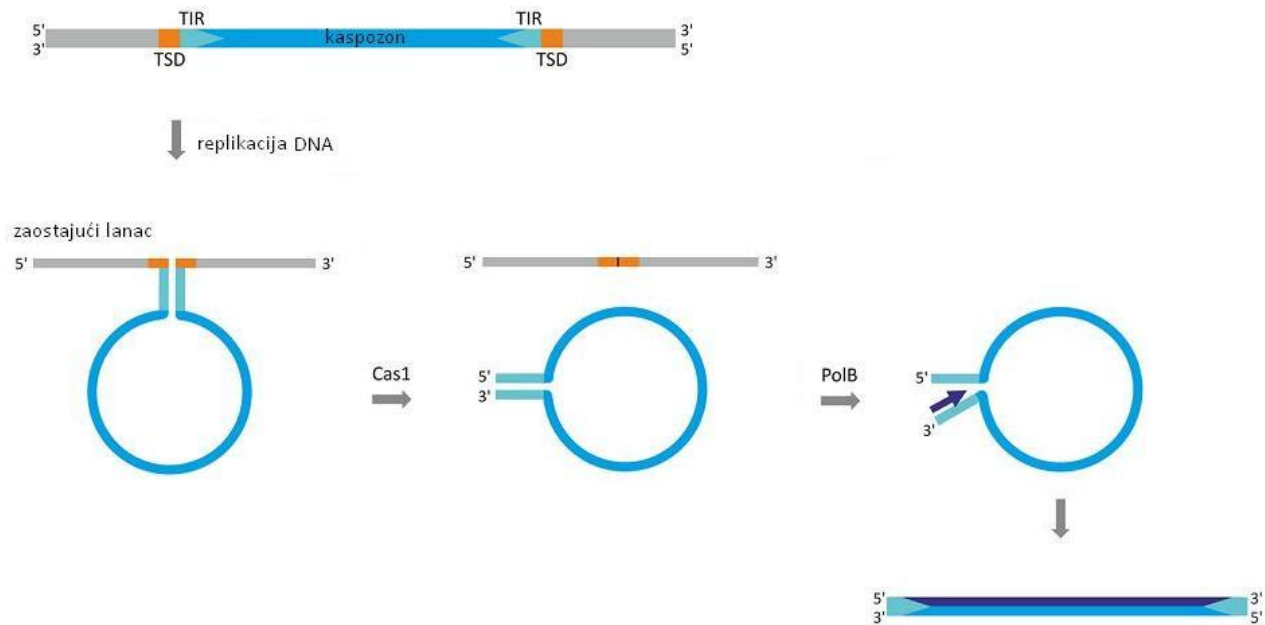
Kaspozoni pokazuju svojstvo ugradnje na specifično mjesto u genomu. Od osamnaest istraženih kaspozona, desetak ih se ugradilo u intergenska područja, tri su se ugradila u gen za elongacijski faktor α EF-2, a pet u različite gene za tRNA. Insercije kaspozona u ove gene događaju se u 3' distalnim regijama, stoga one ne uzrokuju inaktivaciju tih gena. Svi proučeni kaspozoni, osim kaspozona kod vrsta *Methanococcoides burtonii* DSM 6242 i *M. psychrophilus* R15 ugrađeni su na samo jedno mjesto u genomu. U ove dvije vrste nađeni su ostaci dodatnih kopija kaspozona, inaktiviranih nonsense mutacijama (kod *M. burtonii*) ili bez invertnih repeticija (kod *M. psychrophilus*) (Krupovic i sur. 2014).

4. MEHANIZAM IZREZIVANJA I INSERCIJE KASPOZONA

Mehanizam izrezivanja i insercije kaspozona pretpostavljen je na temelju njihove sličnosti s eukariotskim samo-sintetizirajućim DNA transpozonomima - Polintonima/Maverickima i na temelju djelovanja proteina Cas1 u sustavu CRISPR-Cas.

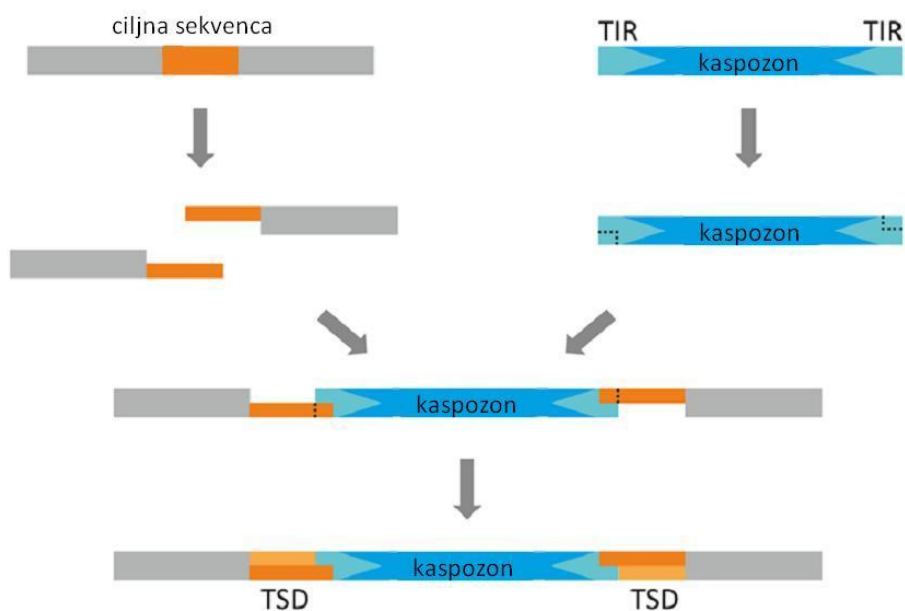
Kaspozonska sekvenca je s obje strane omeđena invertnim repeticijama. Tijekom replikacije DNA, na zaostajećem lancu vjerojatno dolazi do sparivanja invertnih repeticija, te kaspozonska sekvenca formira jednolančanu petlju. Pretpostavlja se da zatim endonukleaza Cas1 reže petlju iz originalnog mjesta u genomu, a sparene invertne repeticije mogu služiti kao ishodište replikacije komplementarnog lanca DNA. Kao početnice mogu poslužiti terminalni proteini vezani na 5'-kraj molekule DNA, ili ih može sintetizirati primaza domaćinske stanice (Krupovic i sur. 2014).

Terminalni proteini se obično vežu na 5' kraj molekule DNA, te interakcijom s DNA polimerazom stvaraju heterodimer. Sinteza novog lanca započinje vezanjem deoskiribonukleotida komplementarnog posljednjem nukleotidu na 3' kraju kalupa na specifičnu hidroksilnu skupinu serina, treonina ili tirozna terminalnog proteina. Ova reakcija katalizirana je DNA polimerazom, te se daljnja sinteza novog lanca nastavlja na uobičajen način (DePamphilis, 1996). Pretpostavlja se da bi se djelovanje terminalnih proteina kodiranih kaspozonomima odvijalo na jednak način, te bi sintezu lanca dalje provodila PolB polimeraza, uz pomoć drugih proteina kodiranih domaćinskim ili kaspozonskim genomima. Postupak izrezivanja kaspozona iz donorske sekvence i replikacije DNA prikazan je na Slici 1.



Slika 1. Predloženi postupak izrezivanja kaspozona (plava boja) iz donorske sekvence i sinteza komplementarnog lanca DNA. Svijetloplavim trokutićima označene su invertne repeticije, a narančastom bojom direktne repeticije. Preuzeto i prilagođeno iz Krupovic i sur. 2014.

Nakon sinteze komplementarnog lanca, kod nekih se kaspozona tupi krajevi vjerojatno dodatno procesiraju na mjestima invertnih repeticija, te nastaju kratki stršeći krajevi djelomično komplementarni ciljnoj sekvenci. Idući korak uključuje stvaranje cik-cak ureza u ciljnoj sekvenci i umetanje kaspozona, vjerojatno pod utjecajem proteina Cas1. Do nastanka direktnih repeticija vjerojatno dolazi popravkom ureza u ciljnoj sekvenci nakon umetanja kaspozona, prikazano na Slici 2 (Krupovic i sur. 2014).



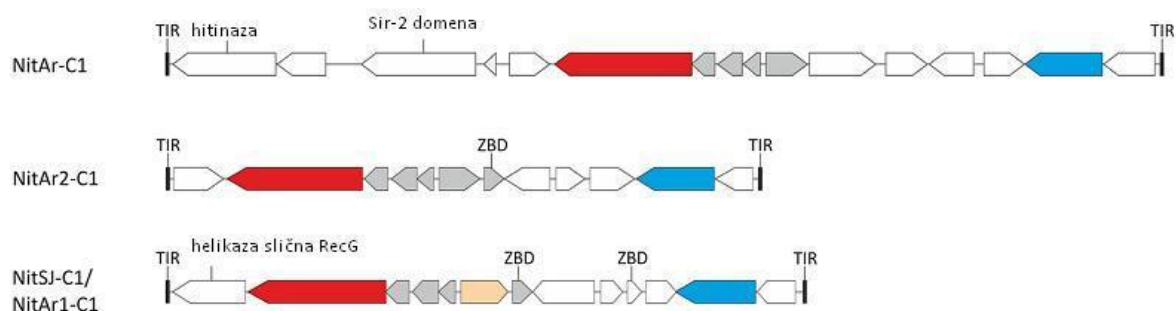
Slika 2. Predloženi postupak insercije kaspozona (plava boja) u ciljnu sekvencu (narančasto). Popravkom ureza u ciljnoj sekvenci nakon insercije kaspozona nastaju direktna ponavljanja (TSD). Preuzeto i prilagođeno iz Krupovic i sur. 2014.

5. SKUPINE KASPOZONA

Podjela kaspozona u tri porodice temelji se na sadržaju njihovih gena, taksonomskim odnosima i razlikama u proteinima Cas1.

5.1 PORODICA 1

U porodicu 1 za sada spadaju 4 kaspozona otkrivena u genomima arheja roda *Nitrosopumilus*, pronađenih u morskim sedimentima. Za ovu porodicu karakterističan je gen za polimerazu PolB kojoj kao početnice služe terminalni proteini vezani na 5' kraj jednolančane DNA. Svi kaspozoni porodice 1 dijele 5 zajedničkih gena: one za proteine Cas1 i PolB, te još tri gena sa zasad neistraženom funkcijom (Slika 3).

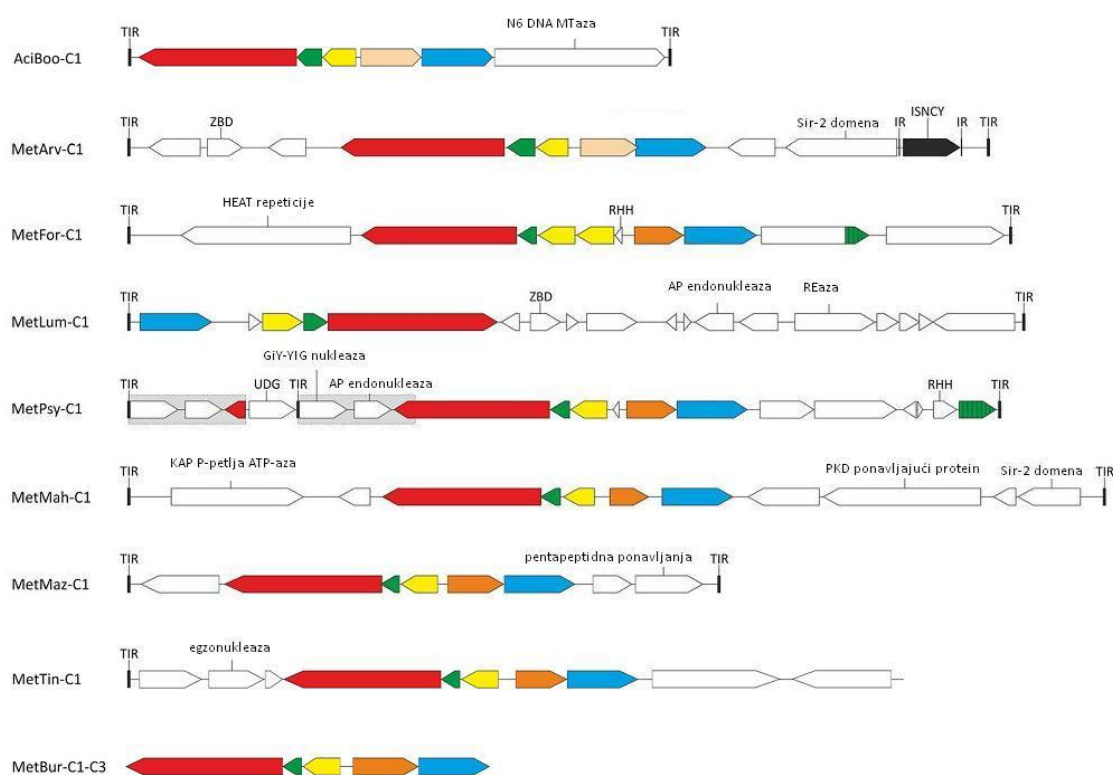


Slika 3. Genomska mapa kaspozona porodice 1. Crvenom bojom označena je DNA PolB, plavom Cas1, ružičastom - neortologna domena sa uzvojnica-okret-uzvojnica domenom; TIR - inverzne repeticije, ZBD - protein sa cink-vezujućom domenom. Preuzeto i prilagođeno iz Krupovic i sur. 2014.

5.2 PORODICA 2

Kaspozoni porodice 2 pronađeni su u različitim arhejama koljena Euriarchaeota. Veličina im varira od 6 do 16 kb, te se razlikuju po sastavu gena. Najčešće sadrže pet zajedničkih gena: za proteine Cas1 i PolB, te za endonukleazu HNH i dva specifična proteina sa uzvojnica-okret-uzvojnica motivom. Proteini Cas1 kaspozona porodice 2 na C-terminusu sadrže specifičnu fuziju s domenom uzvojnica-okret-uzvojnica. Ovakve fuzije ne javljaju se kod proteina Cas1 koji pripadaju kaspozonomima drugih porodica.

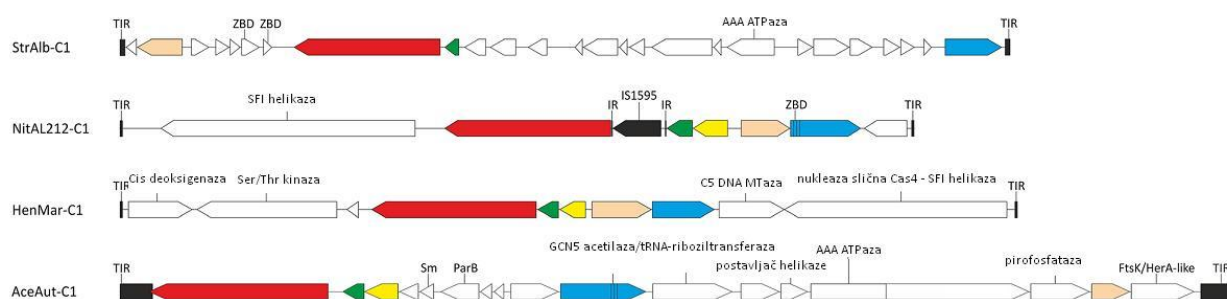
DNA polimeraza PolB kaspozona porodice 2 također sadrži sličnu uzvojnica-okret-uzvojnica domenu na svom C-terminusu. Srodna je porodici arhejskih polimeraza PolB3, koje za stvaranje novog lanca DNA koriste RNA početnice. Jedan od specifičnih proteina sa uzvojnica-okret-uzvojnica motivom sadrži C-terminalnu repetitivnu domenu HEAT - zavojitu strukturu sličnu užetu, uključenu u ostvarivanje interakcija među različitim proteinima. Smatra se da su endonukleaza HNH i specifični proteini sa uzvojnica-okret-uzvojnica motivima uključeni u prepoznavanju ciljnog mjesta, te da sa Cas1 surađuju u integraciji kaspozona u ciljnu sekvencu. Genomska mapa kaspozona porodice 2 prikazana je na Slici 4.



Slika 4. Genomska mapa kaspozona porodice 2. Crvenom bojom označena je DNA PolB, plavom Cas1, zelenom endonukleaza HNH, narančastom domena HEAT, žutom uzvojnica-okret-uzvojnica domena, ružičastom neortologna domena sa uzvojnica-okret-uzvojnica domenom; TIR - inverzne repeticije, ZBD - protein sa cink-vezujućom domenom, MTaza - metiltransferaza, REaza - restriksijska endonukleaza, AP - apurinsko/apirimidinsko, UDG - uracil-DNA glikozilaza, IR - inverzne repeticije. Preuzeto i prilagođeno iz Krupovic i sur. 2014.

5.3 PORODICA 3

Kaspozoni porodice 3 sastoje se od četiri grupe, nađeni su kod različitih vrsta bakterija i pokazuju veliku raznolikost. Protein Cas1 kaspozona NitAL212-C1 sadrži cink-vezujuću domenu vezanu na N-terminus i domenu sa uzvojnica-okret-uzvojnica motivom vezanu na C-terminus. S druge strane, protein Cas1 kaspozona AceAut-C1 sadrži obje domene vezane na C-terminus, a protein Cas1 kaspozona StrAlb-C1 uopće ne sadrži ove dodatne domene. Neki kaspozoni porodice 3 sadrže gene za endonukleazu HNH i očuvani protein s motivom uzvojnica-okret-uzvojnica, zajedničke s kaspozonom porodice 2 (Krupovic i sur. 2014). Genomska mapa kaspozona porodice 3 prikazana je na Slici 5.



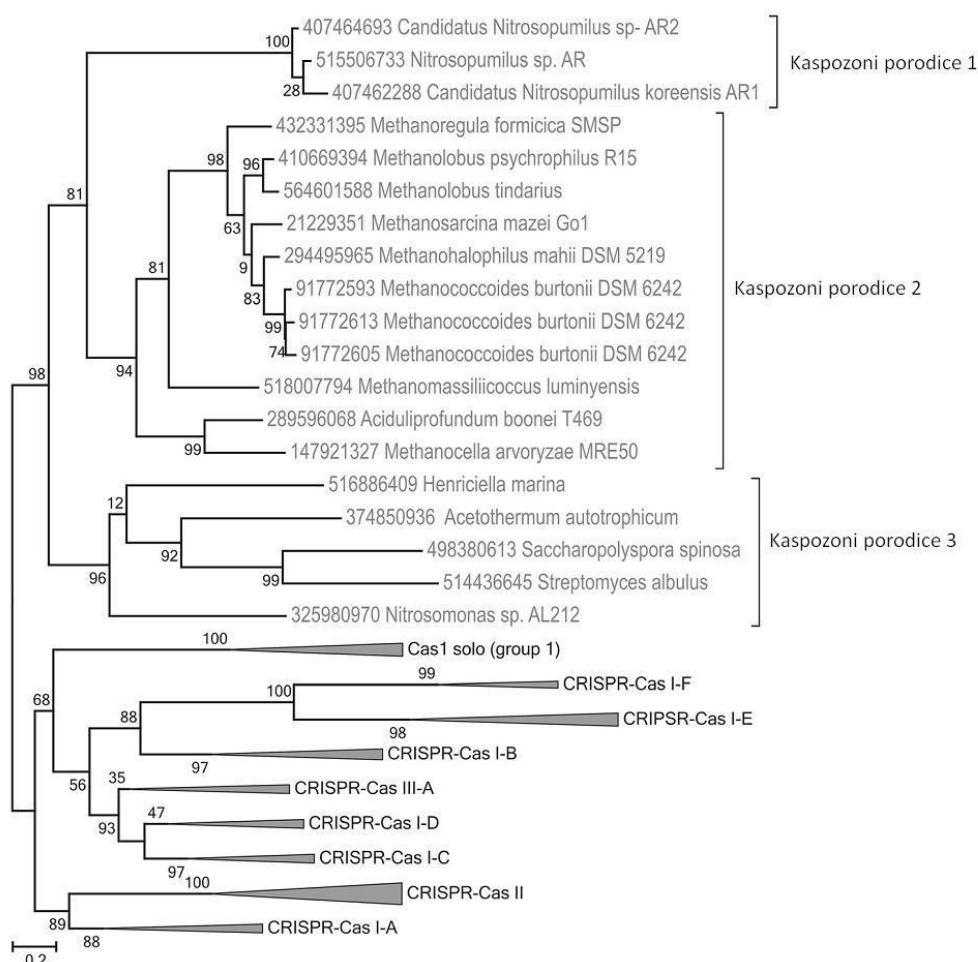
Slika 5. Genomska mapa kaspozona porodice 3. Crvenom bojom je označena DNA PolB, plavom Cas1, zelenom endonukleaza HNH, žutom uzvojnica-okret-uzvojnica domena, ružičastom neortologna domena sa uzvojnica-okret-uzvojnica domenom; TIR - inverzne repeticije, ZBD - protein sa cink-vezujućom domenom, MTaza - metiltransferaza, SFI - superporodica I. Preuzeto i prilagođeno iz Krupovic i sur. 2014.

6. SLIČNOST KASPOZONA I DRUGIH POKRETNIH GENETIČKIH ELEMENATA

Zaključci o srodstvenim odnosima i sličnosti kaspozona i drugih pokretnih genetičkih elemenata zasnovani su prvenstveno na filogenetičkim usporedbama kaspozonskih gena s genima velikog broja virusa, plazmida i transpozona.

Pošto se kaspozoni u većem broju slučajeva javljaju kod arheja, te s obzirom na međusobnu sličnost PolB proteina kaspozona porodica 2 i 3 i arhejskih PolB3 proteina, smatra se da su se kaspozoni razvili kod arheja, te se horizontalnim prijenosom gena prenijeli u bakterije.

Za razliku od svih ostalih do sada poznatih pokretnih genetičkih elemenata, kaspozoni za izrezivanje i inserciju u DNA domaćina koriste isključivo proteine Cas1. Drugi pokretni genetički elementi se za iste radnje oslanjaju na integraze (čije se djelovanje zasniva na katalitičkom središtu s tirozinskim ili serinskim bočnim ogrankom) ili transpozaze (najčešće transpozaze porodice DDE, koje u katalitičkom središtu sadrže dva aspartata i glutamat), međutim, kod kaspozona nije uočeno postojanje gena za ove enzime. Proteini Cas1 u svom katalitičkom središtu sadrže dva aspartata na pozicijama 218 i 221, histidin na poziciji 208 i glutamat na poziciji 141. Također, riješene strukture proteina Cas1 ne ukazuju na strukturnu sličnost sa do sada poznatim transpozazama. Proteini Cas1 kaspozona iz porodice 1 u usporedbi s proteinima Cas1 porodica 2 i 3 čine odvojenu kompaktnu filogenetičku skupinu (Slika 6).



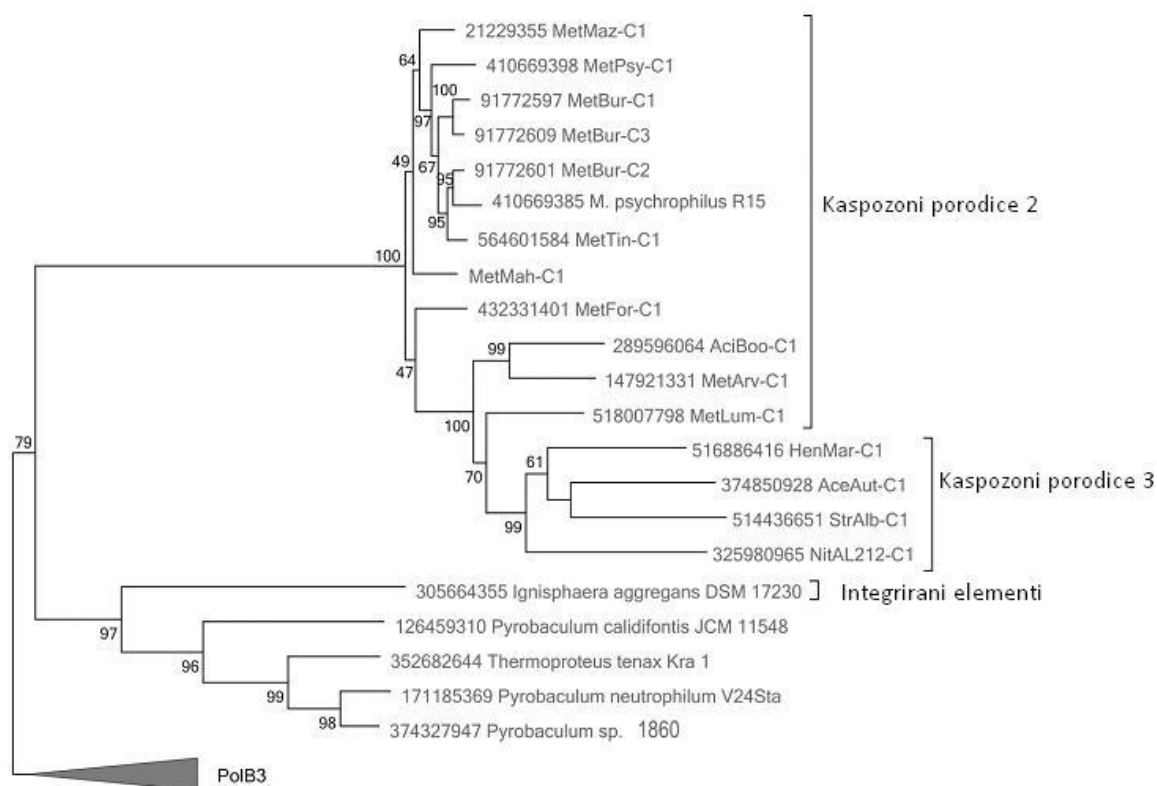
Slika 6. Filogenija Cas1 proteina. Preuzeto i prilagođeno iz Krupovic i sur. 2014.

Kaspozoni i Polintoni/Mavericki spadaju u skupinu samo-sintetizirajućih pokretnih genetičkih elemenata, slične su veličine, te sadrže invertne repeticije. Također i kaspozoni porodice 1 i Polintoni/Mavericki kodiraju srodne DNA polimeraze PolB koje kao početnice za početak replikacije koriste terminalne proteine. Međutim, za razliku od Polintona/Mavericka, kaspozoni ne sadrže gene za kapsidne proteine. Provedene filogenetske analize pokazale su kako ova dva tipa pokretnih genetičkih elemenata međusobno nisu u srodstvu.

PolB polimeraze kaspozona porodice 1 kao početnicu koriste terminalne proteine, te su srodne PolB proteinima arhejskih virusa His1 i His2. Geni za ove proteine srodni su sličnim genima drugih prokariotskih pokretnih genetičkih elemenata. Pošto virusi His1 i His2 spadaju u

različite porodice, pretpostavlja se da su njihovi preci mogli neovisno preuzeti PolB gene od kaspozona porodice 1, iako je za potvrdu ovih pretpostavki potrebno provesti daljnja istraživanja. PolB polimeraze kaspozona porodice 2 srodne su polimerazama porodice PolB3 arhejskog reda *Thermoproteales*, koje kao početnicu koriste RNA. Zbog toga se pretpostavlja da za sintezu početnica kaspozoni porodice 2 koriste primaze domaćinske stanice.

Skupina PolB polimeraza kaspozona porodice 3 nalazi se unutar skupine PolB polimeraza kaspozona porodice 2 (Slika 7).



Slika 7. Filogenija PolB proteina kaspozona porodica 2 i 3. Preuzeto i prilagođeno iz Krupovic i sur. 2014.

Kaspozoni se ugrađuju na mjesno specifičan način, poput eukariotskih transpozona DADA porodice, koji se ugrađuju u gene za tRNA i sRNA. Za razliku od ovih transpozona, koji se ugrađuju u sredinu ciljnih gena i tako ih inaktiviraju, kaspozoni se, poput bakterijskog transpozona Tn7 ugrađuju u nekodirajuće 3'-distalne regije gena, te geni ostaju aktivni.

Poput transpozona, kaspozoni su omeđeni direktnim i invertnim repeticijama. Sparene invertne repeticije na jednolančanoj petlji služe kao mjesto prepoznavanja i ishodište replikacije za PolB polimerazu kod nekih virusa i plazmida. Također kod nekih plazmida početnicu sintetizira terminalni protein kodiran genomom plazmida. Ovako vjerojatno funkcioniraju kaspozoni porodice 1.

U dva kaspozona (MetArv-C1 i NitAL212-C1) pronađene su insercijske sekvence iz porodica ISNCY i IS1595, koje su sadržavale gene za transpozaze, te invertne i direktne repeticije. Ovo dovodi do zaključka da kaspozoni, poput virusa, horizontalnim prijenosom gena mogu steći nove gene. Gen za endonukleazu PolB kaspozona porodice 2, pronađen u genomu bakterije *Ignisphaera aggregans* DSM 17230, umetnut je u otprije postojeći pokretni genetički element koji nosi gene za tirozinske integraze i nije srodan kaspozonima. Ovo upućuje na mogućnost horizontalnog prijenosa gena između kaspozona porodice 2 i ostalih pokretnih genetičkih elemenata (Krupovic i sur. 2014).

7. EVOLUCIJSKI ZNAČAJ PROTEINA Cas1

Protein Cas1 je ključan enzim sustava CRISPR-Cas, uključenog u stvaranje stečene imunosti kod prokariota (Krupovic i sur. 2014).

Sustav CRISPR-Cas je široko rasprostranjen u prokariotskom svijetu i postoji kod približno 40% bakterija, te preko 90% arheja (Horvath, Barrangou, 2010). Regija CRISPR u genomu prokariota sastoji se od kratkih ponavljajućih palindromskih sekvenci, međusobno pravilno razdvojenih razmaknicama koje sadrže stranu DNA. U blizini regije CRISPR nalaze se geni za proteine Cas. U nekim slučajevima, nakon ulaska strane DNA u prokariotsku stanicu, na određenim mjestima dolazi do njenog cijepanja pod utjecajem proteina Cas1 i Cas2 (Horvath, Barrangou, 2010). Dobiveni fragmenti se zatim pomoću istih enzima ugrađuju u regiju CRISPR (Krupovic i sur. 2014), te tako nastaju razmaknice. Transkripcijom regije CRISPR dobiva se dugačka pre-crRNA (pre-CRISPR RNA), koja se zatim pomoću proteina Cas ili kompleksa Cascade (multimera koji se sastoji od više Cas proteina) cijepa na manje fragmente. Ovi fragmenti čine zrele crRNA, a sadrže razmaknicu i kratka djelomična ponavljanja na svojim krajevima. Ponovnim ulaskom strane DNA u stanicu, dolazi do sparivanja crRNA s homolognim sekvencama te DNA. Ovo vezanje je signal koji navodi kompleks Cascade i enzime Cas na cijepanje strane DNA (Horvath, Barrangou, 2010).

S obzirom na filogeniju i svojstva kaspozonskih proteina Cas1 u usporedbi s ostalim proteinima Cas1, pretpostavlja se da su se proteini Cas1 mogli razviti najprije kao komponenta pokretnih genetičkih elemenata. Kasnijom evolucijom moglo je doći do prilagodbe proteina Cas1 iz druge skupine Cas1-solo za sudjelovanje u mehanizmima stečene imunosti kod arheja, te stvaranja sustava CRISPR-Cas. Sustav CRISPR-Cas se kasnije mogao proširiti u bakterijske stanice horizontalnim prijenosom gena. Također se smatra da je i Cas9, ključan protein CRISPR-Cas sustava tipa II, potekao iz transpozonskog proteina. Na ovaj način, kaspozoni bi mogli imati ključnu ulogu u procesu nastanka specifične imunosti kod prokariota, te tako djelovati kao snažni pokretači evolucije. Smatra se da je razvoj specifične imunosti kod eukariota također potekao od pokretnih genetičkih elemenata. Naime, protein RAG1, ključan za procese rearanžmana gena za imunoglobuline kod sisavaca, također se razvio iz DDE transpozaze nađene kod *Transib* transpozona (Krupovic i sur. 2014).

8. ZAKLJUČAK

Kaspozoni čine novo otkrivenu, za sad jedinu poznatu porodicu prokariotskih samo-sintetizirajućih pokretnih genetičkih elemenata. Otkriveni su tijekom filogenetske analize Cas skupine gena, te su za sada dokazani u 19 vrsta arheja i bakterija.

Iako dijele mnoge značajke s različitim drugim prokarotiskim pokretnim genetičkim elementima, kao i eukariotskim transpozonomima skupine Polintona/Mavericka, jedinstveni su po tome što za izrezivanje i integraciju u genom domaćina koriste endonukleazu Cas1. Cas1 je do sada bio povezan prvenstveno sa CRISPR-Cas sustavom stečene imunosti kod prokariota. Filogenetičke analize pokazale su da su kaspozonski proteini Cas1 srodni proteinima Cas1 sustava CRISPR-Cas, stoga ovo otkriće ima velik značaj za objašnjenje nastanka prokariotskog sustava stečene imunosti. Također, ovo otkriće ukazuje na veliku važnost pokretnih genetičkih elemenata kao pokretača evolucije.

Djelovanje proteina Cas1 sustava CRISPR-Cas dokazano je na većem broju supstrata, te se smatra da proteini Cas1 kaspozona funkcioniraju na jednak način. Međutim, do sada nisu napravljeni pokusi koji bi in vivo dokazali sposobnost samoizrezivanja i ponovne insercije kaspozona u genom domaćinske stanice. Još uvijek se moraju otkriti i potvrditi funkcije nekih proteina kodiranih kaspozonomima. Također, kaspozoni su dokazani u samo 19 vrsta arheja i bakterija, te bi otkrića kaspozona u većem broju vrsta pridonijelo njihovom boljem razumijevanju.

Otkriće potpuno novog tipa pokretnih genetičkih elemenata koji su sudjelovali u vrlo važnom evolucijskom procesu (nastanku stečene imunosti kod prokariota) pokazuje koliko su filogenetske metode moćan i snažan biološki alat, te da genomi još uvijek sakrivaju veliki broj neistraženih područja.

9. LITERATURA

- DePamphilis M. L., 1996. DNA Replication in Eukaryotic Cells. Cold Spring Harbor Laboratory Press, New York, 131-177.
- Horvath P., Barrangou R., 2010. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science* **327**, 167-170.
- Krupovic M., Makarova K. S., Forterre P., Prangishvili D., Koonin E. V., 2014. Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biology* **12**, 36-48
- Leplae R., Hebrant A., Wodak S. J., Toussaint A., 2004. ACLAME: A CLAssification of Mobile genetic Elements. *Nucleic Acids Research* **32**, 45-49.
- Lodish H., Berk A., Zipursky S. L., 2000. Molecular Cell Biology. W. H. Freeman, New York, 5. poglavlje dostupno na <http://www.ncbi.nlm.nih.gov/books/NBK21495/>
- Martinez-Abarca F., Toro N., 2002. Group II introns in the bacterial world. *Molecular Microbiology* **38**, 917-926.
- Tenaillon M. I., Hufford M. B., Gaut B. S., Ross-Ibarra J., 2011. Genome size and Transposable Element Content as Determined by High-Throughput Sequencing in Maize and *Zea luxurians*. *Genome Biol Evol* **3**, 219-229.

10. SAŽETAK

Pokretni genetički elementi su definirane sekvence DNA koje imaju sposobnost ugradnje u genom domaćinskih stanica i promjene svoje genetičke lokacije unutar njega. Javljaju se u svim domenama živih organizama, te se smatra da zbog promjena koje unose u genom (npr. inaktivacije gena usred ugradnje ili dopreme novih gena) pokretni genetički elementi djeluju kao jedni od najbitnijih pokretača evolucije.

Kaspozoni su novootkriveni tip prokariostkih samo-sintetizirajućih pokretnih genetičkih elemenata. Iako dijele sličosti s većim brojem različitih pokretnih genetičkih elemenata prokariota i eukariota, jedini za pokretanje u genomu domaćina koriste endonukleaze Cas1. Osim gena za endonukleazu Cas1, svi kaspozoni nose i gen za DNA polimerazu polB. Filogenetskim analizama ovih gena istraženi su srodstveni odnosi kaspozona i drugih pokretnih genetičkih elemenata. S obzirom na filogeniju kaspozonskih proteina Cas1, pretpostavlja se da su proteini Cas1 sustava CRISPR-Cas, odgovornog za postojanje stečene imunosti kod arheja i bakterija, potekli upravo od kaspozonskih proteina Cas1. Ovo čini kaspozone ključnim dijelom razvoja sustava stečene imunosti kod prokariota.

11. SUMMARY

Mobile genetic elements are defined DNA sequences capable of integrating into a host genome and moving from one location in the genome to the other. They are abundant in all three domains of life. Since they can render genes inactive by insertion inside the coding sequence, or bring novel genes to the host cell, mobile genetic elements are considered to be one of the main drivers of evolution.

Casposons are recently discovered type of self-synthesizing prokaryotic mobile genetic elements. Although they share a number of traits with different types of prokaryotic and eukaryotic mobile genetic elements, they are the only ones to use Cas1 endonucleases for excision and integration into the host genome. Phylogenetic analysis was used to explore genetic relationships between casposons and other mobile genetic elements. Based on Cas1 protein phylogeny, it is proposed that Cas1 proteins from CRISPR-Cas adaptive immunity system in prokaryotes emerged from casposon Cas1 proteins. This suggests that casposons played a key role in the origin of acquired immunity in archaea and bacteria.