

Odbor za primijenjenu genomiku  
Hrvatske akademije znanosti i umjetnosti

1. Marija Rosandić

Marija Rosandić Pilaš, prof.dr.sc. dr.med., internistica,  
gastroenterologinja, pulmologinja, u mirovini; Odbor za bioinformatiku  
i biološku fiziku HAZU

2. Uža znanstvena problematika kojom se bavite

Tijekom radnog staža istraživala sam prekanceroze želuca i debelog  
crijeva, endoskopijske metode za zaustavljanje krvarenja iz probavnog  
sustava, infekciju s *H.pylori*.

Od 2001. godine intenzivno se bavim fundamentalnim istraživanjima na  
području genomike: bazičnim otkrićem metode za detekciju repetitivnih  
sekvenci DNA (Key string algorithm, i Global Repeat Map Algorithm  
(GRM)), istraživanjem repetitivnosti višeg reda u humanom genomu i genomima  
primata.

Od 2003. radim u području genomike. Na tim istraživanjima rade Marija  
Rosandić, Matko Glunčić, Vladimir Paar, Ines Vlahović i drugi.  
Stvorena je nova metoda za identifikaciju i analizu periodičnih i  
superperiodičnih (HOR) struktura u genomskoj sekvenci i ab initio  
izrađen odgovarajući vrlo složen kompjutorski softver GRM. U GRM metodi  
po prvi put je uspjele direktno preslikati simboličku DNA sekvencu u  
frekventnu domenu što omogućuje jedinstveno visokoselektivnu  
identifikaciju globalne repetitivne mape i naročito aproksimativne  
superperiodične (supersimetrijske) strukture. Ključna ideja GRM metode  
je uvođenje kompletnog ansambla K-stringova kao robustnog indikatora  
najšireg spektra za identifikaciju aproksimativnih periodičnosti  
primarnog i viših redova. Pomoću GRM algoritma, u genomima čovjeka i  
drugih viših primata identificirali smo preko pedeset dosad nepoznatih  
novih primarnih aproksimativnih repetitivnosti i repetitivnosti višeg reda, te  
disperziranih repetitivnosti. Usporedbom odgovarajućih repetitivnosti dokazano  
je da je značajna razlika između čovjeka i ostalih primata  
koncentrirana u velikim aproksimativnim primarnim repetitivnostima i super-  
repetitivnostima. Dodatno, pronađeno je da se ove razlike javljaju na dvije  
razine: prvo, na nivou razlike pojedinih baza i drugo, na nivou  
razlike strukturne organizacije osnovnih repetitivnih jedinica. Na  
primjer, pronađeno je da je razlika u bazama kod odgovarajućih  
repetitivnih sekvenci kromosoma Y čovjeka i čimpanze iznenađujuće  
velika, ~70%. Primjenjena na cijeli kromosom, ova razlika je na nivou  
~14%, što značajno nadmašuje dosadašnje procjene od 2-5% razlike između  
genoma čovjeka i čimpanze. Kao primjer strukturne razlike,  
pronađen je u ljudskom genomu novi tip dugodosežnog evolucijskog  
ubrzanja, Human Accelerated HOR Regions (HAHORs), koji  
razotkrivaju naglu evoluciju strukturnih HOR organizacija nakon  
separacije čovjeka i čimpanze. Generalno, pokazano je da su razlike  
između čovjeka i čimpanze puno veće u području tandemnih repetitivnosti,  
posebno kod HOR-ova, nego kod genskih kodirajućih sekvenci. Otkrivena  
je dosad potpuno nepoznata super-repetitivnost (čak 46 kopija) u NBPf genu  
koji je povezan s radom ljudskog mozga, a koja je u potpunosti odsutna  
u ostalih visokih primata. Na osnovu ovih rezultata iznesena je u ovim  
publikacijama hipoteza o mogućoj važnosti HAHOR-a kao komponenti u  
višerazinskoj mreži regulacije genske ekspresije.

Od 2012. do 2014. istraživanje je rezultiralo novom klasifikacijom trinukleotida kao osnovnih struktura svakog genoma bez obzira na organizam, otkrićem regulatorne uloge start/stop kodonu sličnih trinukleotida (CLTs) i njihovih ekstenziranih oblika u nekodirajućoj DNA, konstrukcijom novog idealnog genetičkog koda, koji se bazira na simetrijama kodona, umjesto na tradicionalnom arteficialnom alfabetskom nizanju. Dodatno, uveli smo novu ideju proširenih start/stop "codon-like" trinukleotida (CLT) u nekodirajućim DNK sekvencama i odgovarajućih proširenih klastera pomoću koje se izvodi nova klasifikacija alfa satelita kod čovjeka i viših primata. Također, pokazuje se da analiza start /stop CLT proširenih klastera omogućuje značajno veći broj kombinacija mogućih proteina u odnosu na standardne tripletne kodove, i samim tim značajno veći informacijski potencijal genomske sekvence, što pridonosi većoj razlici među organizmima. Naš "idealni" genetički kod ukazuje na vodeću ulogu serina kao "lidera" pri stvaranju koda, uvodi evolucijski način kodnog stvaranja, te isključujući do danas najrašireniju Crickovu "frozen accident" hipotezu. Daljnjim istraživanjem smo pokazali da je DNA čvrsta kvartarna struktura, a ne binarna kao što se do sada smatralo. Bazira se na simetrijama nižeg i višeg reda, koje strogo dirigitiraju na koji se način neki oligo ili mononukleotid može inkorporirati u DNA. Iz tih spoznaja uslijedila je kreacija prirodnog zakona o očuvanju DNA. Za to su izneseni i empirijski argumenti analizom genoma raznih organizama.

### 3. 10 najznačajnijih radova

1. M. Rosandić, V. Paar, I. Basar, Key-string segmentation algorithm and higher-order repeat 16mer (54 copies) in human alpha satellite DNA in chromosome 7, *J. Theor. Biology (Cambridge)* 221 (2003) 29-37.
2. V. Paar, N. Pavin, M. Rosandić, M. Glunčić, I. Basar, R. Pezer, S. Durajlija Žinić, ColorHOR – novel graphical algorithm for fast scan of alpha satellite higher-order repeats and HOR annotation for GenBank sequence of human genome, *Bioinformatics* 21 (2005) 846-852.
3. M. Rosandić, V. Paar, I. Basar, M. Glunčić, N. Pavin, I. Pilaš, CENP-B box and pJ $\alpha$  sequence distribution in human alpha satellite higher-order repeats (HOR), *Chromosome Res.* 14 (2006) 735-753.
4. V. Paar, I. Basar, M. Rosandić, M. Glunčić, Consensus Higher Order Repeats and Frequency of String Distributions in Human Genome, *Curr. Genomics* 8 (2007) 93-111.
5. M. Rosandić, M. Glunčić, V. Paar, I. Basar, The role of alphoid higher order repeats (HORs) in the centromere foldng, *J. Theor. Biol.* 254 (2008) 555-560.
6. V. Paar, N. Pavin, I. Basar, M. Rosandić, M. Glunčić, N. Paar, Hierarchical structure of cascade of primary and secondary periodicities in Fourier power spectrum of alphoid higher order repeats, *BMC Bioinformatics* 9 (2008) 466 (BioMed Central), doi: 10.1186/1471-2105-9-466, 1-14.
7. V. Paar, M. Glunčić, I. Basar, M. Rosandić, P. Paar, M. Cvitković, Large tandem, higher order repeats and regularly dispersed repeat units contribute substantially to divergence between human and chimpanzee Y chromosomes, *J. Mol. Evol.* 72 (2011) 34-55.
8. V. Paar, M. Glunčić, M. Rosandić, I. Basar, I. Vlahović, Intragene higher order repeats in neuroblastoma breakpoint family genes distinguish humans from chimpanzees, *Mol. Biol. Evol.* 28 (2011) 1877-1892.

9. M. Rosandić, M. Glunčić, V. Paar, Start/stop codon like trinucleotides extensions in primate alpha satellites, *J. Theor. Biol.* 317 (2013) 301-309.
10. M. Rosandić, V. Paar, M. Glunčić, Fundamental role of start/stop regulators in whole DNA and new trinucleotide classification, *Gene* 531, 2 (2013) 184-190.
11. Marija Rosandić, Vladimir Paar, Codon sextets with leading role of serine create "ideal" symmetry classification scheme of the genetic code, *Gene* 543 (2014) 45-52.

4. Znanstveno-istraživački projekti koje ste vodili ili u njima surađivali.

Voditelj tri domaća znanstvena projekta iz gastroenterologije "Tkivni karcinoembrionalni antigen u probavnom sustavu (1986.-1998.)".

Suradnik na domaćem projektu:

KSA kompjutorsko istraživanje repeticija višeg reda i duplikona u humanom genomu. Istraživanja tumorskih bolesti. (119-09822464-1253) (2007.-2012.).

5. Znanstveno-istraživački projekt koji vodite ili u njemu surađujete.

HRZZ-u je prijavljen prijedlog projekta.

6. Međunarodna suradnja.

Korištenje sekvencioniranih genoma s otvorene NCBI internet adrese.

Korištenje sekvencioniranog genoma neandertalca koji smo dobili od znanstvenika iz Max Planck instituta za humanu genetiku iz Dresdena. Naše teorijsko istraživanje je fundamentalno i s ciljem razumijevanja osnovnih prirodnih zakona u genomici, te tako od općeg značenja, i predstavlja velikim dijelom razvoj vlastitih novih ideja. Istraživanje je individualno i nije zahtijevalo široki krug znanstvene suradnje, nego se suradnja odvija u okviru domaće grupe (Vladimir Paar, Marija Rosandić, Matko Glunčić, Ines Vlahović), uključujući snažnu kompjutorsku podršku.

Sav kompjutorski softver za istraživanja je naš vlastiti, napravljen koristeći tri nove matematičke metode koje smo sami otkrili i razvili. Budućim širenjem naših istraživanja na karakterizaciju nekih bolesti genomskim superrepeticijama planiramo proširenje suradnje na inozemne znanstvenike.

7. Najvažnija oprema Vašeg laboratorija  
Stolno i prijenosno računalo.

9. Oprema koju bi valjalo nabaviti

- na razini laboratorija

Dodatna računalna oprema za studente i doktorande koji surađuju na istraživanjima.

- na razini ustanova

Klaster računala za automatsku pretragu i usporedbu (upotrebom GRM računalnog programa) repeticija i proširenih trinukleotida unutar genomskih sekvenci čovjeka i ostalih eukariota. Serversko računalo za online bazu (na otvorenom web portalu) repeticija u genomima svih eukariota. Nakon izrade kompletne baze jedan smjer istraživanja vodio bi na izradu evolucijske slike vremenskog i prostornog rasporeda svih mutacija unutar odgovarajućih repeticija. Na osnovi ovako dobivene slike bilo bi moguće izraditi i pojednostavljeni fizikalni model evolucije repetitivnih jedinica od genoma nižih primata do genoma čovjeka. Drugi smjer istraživanja modelirao bi razvoj repetitivnih

struktura višeg reda, s posebnim naglaskom na područja nagle akceleracije (fazni prijelazi). Ova dva modela pomogla bi u razumijevanju same evolucije kompletne genske sekvence, organizacije genoma, te uloge nekodirajućeg dijela genoma u regulacijskoj mreži ekspresije gena.

10. Vaše mišljenje o izobrazbi iz genomike u Vašem području, sadašnje stanje i kako unaprijediti:

- dodiplomska
- poslijediplomska doktorska
- poslijediplomska specijalistička

Sadašnje stanje može se najbolje unaprijediti uključivanjem doktorskih studenata i postdoktoranda u naša istraživanja.

11. Sažeto prikažite primjenu genomike u Vašem području u nas i usporedite sa svijetom.

Primjene rezultata naših istraživanja trenutno kod nas ne postoji. U svijetu se sve više pokušava mutacije u nekodirajućem dijelu genoma, posebno unutar repetitivnih jedinica, povezati s regulacijom ekspresije gena i pronaći mehanizme njihovog djelovanja i aktivacije. U svrhu toga objavljuju se algoritmi za pretragu i uspostavljaju se baze repeticija i razlika unutar nekodirajućeg dijela genoma. Ono što mi u ovom trenutku imamo je jedinstven u svijetu brz i robustan algoritam i odgovarajući vlastiti kompjutorski program GRM za automatsku identifikaciju i analizu svih repeticija u bilo kojem genomu. Također, koncepti ekstenziranih "codon like trinukleotida" kao i nove klasifikacije genetskog koda i njegove evolucije, te otkriće mehanizma kvadrupleta su naše originalne ideje i metode.

12. Da li su u nas zadovoljavajući etički i pravni vidovi primjene genomike? Sažeto prikažite u Vašem području i navedite prijedloge. Zbog javne dostupnosti genomskih sekvenci i teorijskog karaktera istraživanja nema etičkih problema za ova istraživanja.

13. Da li je u nas dovoljno razvijena informatička potpora primjeni genomike? Što predlažete?

Vidi točku 9

14. Da li su Vam dostupne sve genomičke baze podataka, ako nisu što predlažete da se poduzme.

Sve genomske baze koje koristimo su otvorenog tipa, uz mogućnost dobivanja dodatnih sekvenci na upit.

15. Kako unaprijediti suradnju u genomici u nas (primjerice ustroj genomičke informacijske mreže)?

Uspostavom jedinstvene baze svih repeticija, te otvorene računalne mreže koja bi omogućila pretragu i usporedbu rezultata s već poznatim repeticijama pohranjenim u bazi, bilo bi moguće povezati mutacije unutar repeticija s kliničkim odstupanjima za svakog pojedinačnog pacijenta. Ova primjena i paralelno istraživanje dovelo bi do otkrića uloge nekodirajućeg dijela genoma u regulacijskoj mreži ekspresije gena i njihovog utjecaja na razvoj pojedinih bolesti. Uspostava mreže s različitim kliničkim centrima dovela bi doznatnog razvoja personalizirane i prediktivne medicine. Na ovaj način postigao bi se

napredak u dijagnosticiranju bolesti koje nisu povezane s monogenetskim poremećajima i omogućio dovoljno rani terapijski tretman.