

Plan upravljanja istraživačkim podacima za HRZZ projekt IP-2022-10-7874 : Alternativno prekrajanje gena BPM2 kao mehanizam uspostave funkcionalne raznolikosti porodice proteina MATH-BTB u uročnjaku

...

Bauer, Nataša

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2024**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:389862>

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

Download date / Datum preuzimanja: **2025-01-14**



Repository / Repozitorij:

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



PLAN UPRAVLJANJA ISTRAŽIVAČKIM PODACIMA (PUP)

Opće informacije		
	Ime i prezime predlagatelja	Nataša Bauer
	Matična organizacija	Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet
	Naziv projekta	Alternativno prekrajanje gena <i>BPM2</i> kao mehanizam uspostave funkcionalne raznolikosti porodice proteina MATH-BTB u uročnjaku <i>Arabidopsis thaliana</i>
	Upravitelj podacima	Nataša Bauer
1. Prikupljanje podataka i dokumentacija		
	Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite formate, vrste i opseg svih podataka s kojima ćete raditi, a ne samo krajnji skup podataka koji će biti rezultat istraživanja)	<p>Tijekom projekta generirat ćemo pet različitih vrsta sirovih podataka:</p> <ol style="list-style-type: none"> 1. sirovi podaci s real-time pcr 2. mikroskopske slike 3. nukleotidne sekvence (Service) 4. proteinske sekvence (generirane na temelju nukleotidnih) 5. sirovi podaci spektrometrije masa peptida <p>Svi podaci biti će pohranjeni u digitalnom obliku u formatu koji se dobije izravno s instrumenata (primjerice, <i>TIFF</i> obliku datoteka za slike dobivene fluorescencijskim mikroskopom; <i>micRUN</i> datoteke za podatke real-time pcr MIC PCR; <i>FASTAQ</i> i <i>FASTAP</i> obliku za nukleotidne i proteinske sekvence. Slike dobivene svjetlosnim mikroskopom će se sortirati i samo odabrane slike bit će spremljene u originalnom formatu specifičnom za svaki korišteni mikroskop. Sirovi podaci spektrometrije masa biti će pohranjeni na lokalnom računalu, na OneDrive ili Puh (https://www.srce.unizg.hr/puh) te po objavi dostupni na ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org).</p> <p>Prikupljati će se i tekstualni podaci generirani izvješćima, člancima, prezentacijama, u docx i pdf formatu. Svi podaci će biti pohranjeni u zaštićenim formatima dobiveni na instrumentima.</p> <p>Mjerenja i kvantifikacija slika/podataka (prikazi u obliku tablica) snimit će se u <i>EXCEL</i> obliku (a za dugotrajnu pohranu, konvertirat će se u <i>CSV</i> oblik).</p>
	Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete, načine organiziranja podataka te alate i instrumente kojima ćete se koristiti za prikupljanje i obradu)	<p>Analitički podaci za real time pcr prikupljaju se s instrumenta koji ih generira (MIC PCR) i obrađuje u matičnom programu te će rezultati po obradi biti pohranjeni u obliku Excel tablica.</p> <p>Slike s fluorescencijskog mikroskopa pohranjivati će se u datoteke, grupirane po tipu uzorka i kronološki, na nekom od virtualnih poslužitelja (OneDrive ili slično). Obrada slika provodi se u aplikacijama sa slobodnim pristupom kao što je ImageJ ili u programu Adobe Photoshop. Skup slika/podataka pratit će i Excel dokumenti s opisom uzoraka. Također, biti će zapisana i odstupanja od protokola i ostale korisne informacije.</p> <p>Sirovi podaci o nukleotidnim i proteinskim sljedovima pohranjivati će se lokalno na računalima suradnika na projektu, uz dodatno pohranjivanje na nekom od virtualnih poslužitelja (OneDrive ili slično).</p> <p>U svim eksperimentima biti će uključena odgovarajuće kontrole, čime se osigurava valjanost podataka.</p> <p>Dosljednost podatka procijenit će se usporedbom ponovljenih mjerenja.</p>
	Koju ćete dokumentaciju i metapodatke izraditi osim podataka? (dokumentacija mora sadržavati informacije i standarde potrebne korisnicima kako bi mogli samostalno	<p>Podaci će biti popraćeni detaljnim opisom svakog uzorka i eksperimentalnog postupka.</p> <p>Opisi će se pohranjivati u obliku <i>README.txt</i> datoteka koje će biti dostupne svim suradnicima na projektu</p>

	čitati i interpretirati podatke u budućnosti, primjerice, kodne knjige, <i>ReadMe</i> datoteke i sl.)	
2.	Pravna i sigurnosna pitanja	
	Jeste li ograničeni sporazumom o povjerljivosti? Imate li potrebna dopuštenja za prikupljanje, obradu, čuvanje i dijeljenje podataka? Jesu li osobe čiji se podaci obrađuju informirani o tome i jesu li dali privolu? Kojim ćete se metodama koristiti u svrhu zaštite osjetljivih podataka (GDPR - posebne kategorije osobnih podataka, navesti metode anonimizacije podataka)?	Pri izvedbi projekta neće se kršiti etička načela. Projekt se provodi na biljkama isključivo u kontroliranim eksperimentalnim uvjetima, u Zavodu za molekularnu biologiju PMF-a. Sporazum o povjerljivosti nije primjenjiv.
	Kako će se regulirati pristup podacima i njihova sigurnost? Koji su potencijalni rizici koje treba uzeti u obzir? Kako ćete osigurati sigurnost pohrane osjetljivih podataka?	Podaci će se obrađivati i njima upravljati u zaštićenom nemrežnom okruženju koristeći se virtualnom desktop tehnologijom. Dio podataka bit će pohranjen u sustavu OneDrive. Pristup podacima tijekom trajanja projekta imaju članovi istraživačkog tima. Pristup sirovim podacima koji se koriste za dobivanje rezultata biti će omogućen istraživačkoj grupi ili po objavi znanstvenih publikacija u odabranom repozitoriju. Pristup podacima i rezultatima objavljenim u znanstvenim publikacijama biti će javno dostupan s obzirom na standarde u polju istraživanja te putem nekih od javnih baza podataka za istraživanje (npr. NCBI, TAIR, ProteomeXchange Consortium). Laboratorijski dnevnicu u digitalnom obliku čuvati će se na računalima suradnika, a fizički laboratorijski dnevnicu rada u laboratoriju, te će biti dostupni glavnom istraživaču. Dodatni elektronički podaci, obrađeni rezultati i prikazi pohranit će se na računalu glavnog istraživača, koji redovito izrađuje sigurnosne kopije. Trajna pohrana velikih setova biti će izvršena na vanjske diskove u najmanje dvije kopije kako bi se zaštitio integritet podataka.
	Kako ćete upravljati zaštitom autorskih prava i drugog intelektualnog vlasništva? Tko će biti vlasnik podataka? Koje će se licencije primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka?	Ne očekuje se da će rezultat istraživanja dovesti do patenta. Ostali problemi intelektualnog vlasništva će se u rješavati prema preporukama Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu. Budući da podaci nisu podvrgnuti ugovoru, te se neće patentirati, objavit će se kao otvoreni podaci pod licencijom <i>Creative Commons</i> CCO po preporuci Europske komisije kako bi omogućili ponovno i ispravno korištenje podataka od strane drugih znanstvenika.
3.	Pohrana i čuvanje podataka	

<p>Kako će radne verzije podataka biti pohranjene tijekom projekta? Kako će se napraviti sigurnosne kopije tih podataka (<i>backup</i>)? Koja je očekivana količina podataka koja će se prikupiti i čuvati tijekom projekta (izraženo u MB/GB/TB)?</p>	<p>Jedna kopija radne verzije podataka biti će pohranjena na lokalnim računalima kako bi se omogućio brz pristup i manipulacija s podacima. Podaci će se tijekom istraživanja s računala glavnog istraživača kopirati u nacionalni sustav za pohranu i dijeljenje podataka Puh (https://www.srce.unizg.hr/puh) koji članovima projektnog tima omogućava pristup aktualnoj verziji podataka i na kojem se dnevno automatizirano izrađuje sigurnosna kopija podataka. Najvažniji podaci biti će čuvani u najmanje dvije kopije kako bi se mogli vjerno replicirati rezultati istraživanja.</p>
<p>Kako će se završne verzije podataka dugotrajno pohraniti i čuvati (i nakon završetka projekta)? U kojim će se formatima čuvati podaci? Koja je očekivana količina podataka koja će se trajno pohraniti (izraženo u MB/GB/TB)?</p>	<p>Podatke ćemo trajno čuvati na virtualnom poslužitelju OneDrive ili PUH (Srce) i također pohraniti u odgovarajući arhiv za podatke na kraju projekta. Tablične podatke čuvat ćemo u CSV obliku, a tekstualne u DOCX (Office Open XML) te PDF-A obliku. DOC oblik će se konvertirati u DOCX oblik. FASTQ datoteke biti će po potrebi komprimirane putem GNU zip programa i pohranjene u GZ obliku. Očekivana količina podataka koja će se trajno pohraniti je oko 300 GB.</p>
<p>4. Dijeljenje i ponovna uporaba podataka</p>	
<p>Kako i gdje će se podaci dijeliti? Koji repozitorij će se koristiti za dijeljenje podataka? Kako će potencijalni korisnici doznati za podatke?</p>	<p>Podaci o sekvencama biti će podijeljeni na javnim repozitorijima poput NCBI ili TAIR. Podaci o interakcijama biti će podijeljeni na ProteomeXchange Consortium. Korisnici će o postojanju tih podataka biti informirani u sklopu objavljenih publikacija.</p>
<p>Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.</p>	<p>Podaci neophodni za bilo koju publikaciju bit će dostupni u trenutku objavljivanja. Svi neobjavljeni podaci pohranit će se u repozitoriju tri godine od završetka projekta.</p>
<p>Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima FAIR-a.</p>	<p>DA Digitalni repozitorij koji će biti korišten podržava načela FAIR-a</p>
<p>Potvrdite da ćete se koristiti digitalnim repozitorijem koji održava neprofitna organizacija (ako ne, objasnite zašto ne možete dijeliti podatke na digitalnom repozitoriju koji nije komercijalan).</p>	<p>Potvrđujem da ćemo koristiti digitalni repozitorij Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu uspostavljenom na sustavu Dabar koji održava PMF kao neprofitna organizacija. Jedna kopija digitalnih podataka će biti pohranjena na javno dostupnim nekomercijalnim bazama Hrvatske znanstvene infrastrukture (PUH).</p>