

НАСТАВНО-НАУЧНОМ ВЕЋУ УНИВЕРЗИТЕТА У БЕОГРАДУ - БИОЛОШКОГ ФАКУЛТЕТА

На VII редовној седници Наставно-научног већа Биолошког факултета Универзитета у Београду, одржаној 08.04.2022. године, на основу молбе ментора, др Милана Драгићевића, вишег научног сарадника Института за биолошка истраживања „Синиша Станковић“, Института од националног значаја за Републику Србију, Универзитета у Београду и др Иване Драгићевић, ванредног професора Биолошког факултета, Универзитета у Београду, одређена је Комисија за преглед и оцену докторске дисертације **Данијеле М. Пауновић** (рођене Јонтуловић), истраживача сарадника Универзитета у Београду, Института за биолошка истраживања „Синиша Станковић“, Института од националног значаја за Републику Србију, Универзитета у Београду под насловом: **„Идентификација *AGP* гена кичице (*Centaurium erythraea*, *Gentianaceae*) и праћење њихове експресије у одговору на механичке повреде биљног ткива гајеног *in vitro*“** у саставу: др Тијана Цветић Антић, ванредни професор Биолошког факултета, Универзитета у Београду; др Марко Ђорђевић, ванредни професор Биолошког факултета, Универзитета у Београду и др Марија Марковић, виши научни сарадник, Института за биолошка истраживања „Синиша Станковић“, Института од националног значаја за Републику Србију, Универзитета у Београду.

Комисија је прегледала урађену докторску дисертацију кандидаткиње и Наставно-научном већу Универзитета у Београду - Биолошког факултета подноси следећи

ИЗВЕШТАЈ

Општи подаци о докторској дисертацији

Докторска дисертација Данијеле М. Пауновић, под насловом „Идентификација *AGP* гена кичице (*Centaurium erythraea*, *Gentianaceae*) и праћење њихове експресије у одговору на механичке повреде биљног ткива гајеног *in vitro*“ се састоји од укупно 137 страна, од којих је 117 нумерисано, и састоји се од следећих поглавља: **Увод** (22 стране, стр. 1-22), **Циљеви рада** (1 страна, стр. 23), **Материјал и методе** (23 стране, стр. 24-46), **Резултати** (36 страна, стр. 47-82), **Дискусија** (14 страна, стр. 83-96), **Закључци** (1 страна, стр. 97) и **Литература** (19 страна, стр. 98-116). Поред наведеног, докторска дисертација обухвата и биографију аутора (стр. 117), као и следеће целине: насловну страну на српском и енглеском језику, сажетак дисертације на српском и енглеском језику, списак ментора и чланова комисије, захвалницу, листу скраћеница, садржај, изјаву о ауторству, изјаву о истовестности штампане и електронске верзије докторског рада и изјаву о коришћењу.

Докторска дисертација садржи укупно 30 слика (4 у поглављу Увод, 25 у поглављу Резултати и 1 у поглављу Дискусија) и 20 табела (2 табеле у Уводу, 14 у поглављу Материјал и методе и 4 у поглављу Резултати). Поглавље Литература садржи 273 библиографске јединице.

Анализа докторске дисертације

Докторска дисертација Данијеле Пауновић припада ужој научној области физиологије и молекуларне биологије биљака. Основни циљеви истраживања су идентификација гена за арабиногалактанске протеине (*AGP*) кичице (*Centaurium erythraea*) и испитивање њихове експресије коришћењем различитих експерименталних система како би се формирала слика у којим физиолошким процесима производи ових гена учествују.

Наведени циљеви су остварени кроз успостављање нове методологије за идентификацију *AGP* гена на основу секвенце, базиране на машинском учењу, која је потом искоришћена за идентификацију *AGP* гена кичице. Експресија одабраних *AGP* гена је затим испитана у одговору на механичке повреде биљног ткива гајеног *in vitro*, у узорцима ткива из различитих фаза развића биљака кичице, као и након гајења биљака на растућим концентрацијама *Yariv* реагенса (β GlcY) који специфично везује и преципитира *AGP*.

Поглавље **УВОД** је подељено на три целине у којима је изнет преглед литературе која чини теоријски оквир теме дисертације и омогућава тумачење остварених резултата и разумевање њиховог значаја. Прва целина садржи преглед досадашњег знања о биљним гликопротеинима богатим хидроксипролином (*HRGP*). Предмет истраживања дисертације - арабиногалактански протеини чине поткласу *HRGP*, па је у оквиру ове целине описана подела и биосинтеза *AGP*, уз преглед експериментално утврђених физиолошких улога, са фокусом на њихове улоге у одговору биљног ткива на механичке повреде и у разноврсним процесима током развића биљке. Пошто је један од основних циљева докторске дисертације побољшање начина за идентификацију секвенци *AGP*, описане су методе из литературе, уз образложење њихових недостатака, који се огледају у немогућности идентификације секвенци са ниским садржајем аминокиселина и малом заступљеношћу мотива карактеристичних за овај тип секвенци. У другом потпоглављу су описане опште карактеристике кичице и досадашња истраживања везана за *AGP* ове биљке. Кичицу одликује изванредна развојна пластичност пошто поседује велики потенцијал за регенерацију у условима *in vitro*, а досадашња истраживања *AGP* су била концентрисана на испитивање њихове улоге и динамике током развојних процеса кичице. Ипак, услед непостојања транскриптома кичице са добром покривеношћу и ограничења метода за идентификацију секвенци *AGP*, само је неколико гена за *AGP* кичице идентификовано у претходним истраживањима. На крају увода је дат преглед основних појмова из области машинског учења, укључујући обучавање и процену модела у надгледаном машинском учењу, принципе бинарне класификације и метрика за процену перформанси карактеристичних за овај тип надгледаног машинског учења. Описана су четири алгорита машинског учења коришћена за прављење модела у оквиру дисертације, и дат је осврт на употребу машинског учења за предвиђања на основу протеинске секвенце, са акцентом на предвиђање хидроксилације пролина.

У оквиру поглавља **ЦИЉЕВИ РАДА** дефинисани су циљеви истраживања докторске дисертације који обухватају: успостављање методологије за прецизну идентификацију *AGP* секвенци која ће послужити за идентификацију гена за *AGP* из *de novo* састављеног транскриптома кичице. Како би се *AGP* гени повезали са физиолошким процесима у којима учествују, у одабраном подскупу идентификованих гена испитивана је експресија: 48 сати након механичке повреде листа и корена биљака кичице гајених у условима *in vitro*; у експлантатима листа и корена кичице гајених на различитим концентрацијама *Yariv* реагенса; у различитим узорцима биљака гајених *in vitro*, биљака из природе и из различитих морфогенетских процеса кичице (соматске ембриогенезе и органогенезе).

Поглавље **МАТЕРИЈАЛ И МЕТОДЕ** садржи шест потпоглавља од којих су нека подељена у више одељака. У оквиру првог потпоглавља је представљена методологија успостављања иновативног концепта за филтрирање и анализу *HRGP* секвенци, која се заснива на моделу машинског учења за предвиђање вероватноће хидроксилације пролина у биљним протеинима. Детаљно је објашњен начин тренирања модела, од припреме података

за тренирање и евалуацију, трансформације протеинских секвенци у нумерички облик, преко одабира атрибута коришћених за тренирање, оптимизације хиперпараметара и евалуације модела. Дат је опис и осталих алата који су, у комбинацији са развијеним моделом, искоришћени за идентификацију и анализу HRGP секвенци. Поменути алати чине софтверски пакет *ragp*, написан у R програмском језику, који је креиран у оквиру ове тезе. Потом је описан поступак идентификације HRGP секвенци *C. erythraea* коришћењем *ragp* пакета из *de novo* састављеног транскриптома кичице. Следеће потпоглавље обухвата методологију филогенетске анализе одабраних AGP секвенци кичице базиране на поређењу са хомолозима из 18 одабраних биљних врста.

У четвртном потпоглављу је описано успостављање *in vitro* културе коренова и изданака кичице, дизајн експеримената за испитивање утицаја механичке повреде на експресију AGP гена кичице гајене у условима *in vitro*, као и испитивање ефекта *Yariv* реагенса на експресију AGP гена кичице гајене у условима *in vitro*. Описани су типови узорака и начини узорковања (ткива биљака гајених без регулатора растења или ткива биљака из различитих фаза органогенезе или соматске ембриогенезе) који су коришћени за испитивање експресије AGP гена.

Потпоглавље пет садржи преглед коришћених молекуларно биолошких метода које обухватају методе за изолацију и проверу квалитета РНК, синтезу комплементарне ДНК, дизајнирање прајмера за одабране AGP гене, протоколе за PCR и квантитативни PCR, и поступак одабира референтних гена за нормализацију експресије. На крају је описана статистичка анализа података о генској експресији.

Поглавље **РЕЗУЛТАТИ** подељено је на три потпоглавља у којима су представљени резултати добијени током истраживања уз одговарајуће табеларне и графичке приказе. Легенде слика пружају јасна објашњења представљених резултата.

Прва целина садржи резултате тренирања и евалуације модела за предвиђање вероватноће хидроксилације пролина на основу локалне аминокиселинске секвенце. Од испробаних 16 комбинација алгоритама које су чиниле међусобне комбинације четири алгорита за машинско учење: k-најближих суседа (KNN), метод потпорних вектора (SVM), ансамбли независних стабала одлучивања (RF) и појачавање по градијенту (XGB) и три различита приступа за одабир атрибута: минимална сувишност максимална значајност (mRMR), удео добитка информације (IGr) и секвенциона претрага унапред (sfs), уз приступ без одабира атрибута, показано је да најбоље перформансе постиже комбинација алгорита појачавања по градијенту комбинована са секвенционом претрагом унапред као методом за одабир група атрибута које представљају локалне секвенце протеина у векторском простору. Затим су представљени резултати оптимизације прага одлуке, утицај дужине локалне аминокиселинске секвенце коришћене за обучавање на перформансе модела, као и поређење перформанси добијеног модела са већ доступним серверима за исту намену, на основу кога је јасно да је развијени модел подеснији за предвиђање позиција хидроксипролина у биљним протеинима. Приказ утицајности различитих атрибута на предвиђања модела показао је да мали број коришћених атрибута има утицај на предвиђања модела. На крају ове целине представљени су резултати анотације HRGP секвенци из 62 биљна протеома који су анализирани коришћењем *ragp* пакета. Овом анализом идентификовано је преко 30 хиљада потенцијалних арабиногалактанских секвенци из 62 биљна протеома које су асоциране са различитим конзервираним доменима.

Друго потпоглавље садржи резултате идентификације, анализе и класификације HRGP секвенци *C. erythraea*. Резултати су представљени бројним дијаграмима који

олакшавају праћење токова анализа и дају увид у хетерогеност идентификованих секвенци. Од идентификованих AGP секвенци одабрано је шест секвенци са фасциклинским доменом (FLA), шест са протеин киназним доменом (KLA) и шест арабиногалактанских пептида (AGP) чије су структурне одлике детаљније испитане и чија је експресија даље праћена у различитим експерименталним системима што је представљено у наредном потпоглављу. Анализом експресије одабраних гена у одговору на стрес изазван механичким повредама експлантата листова и коренова показано је да су два арабиногалактанска пептида, индукована након повреде листа, а репримирана након повреде корена. У експерименту у коме је праћена експресија одабраних AGP гена након дуготрајног излагања експлантата различитим концентрацијама *Yariv* реагенса, забележене су мале промене експресије свих испитиваних гена. Анализом експресије одабраних гена у различитим узорцима ткива и органа *in vitro* гајених биљака, ткива из различитих фаза органогенезе и соматске ембриогенезе као и ткива из природе, показано је да већина испитиваних гена има значајне промене експресије барем у неким од испитиваних узорака.

У поглављу **ДИСКУСИЈА** кандидат је уз коришћење литературних извора разматрао претходно изнете резултате. Дискусија је организована на истоветан начин као поглавље резултати, па је у првој целини образложен изабрани поступак тренирања и евалуације развијеног модела у контексту објективне процене перформанси. Поступак је подржан релевантним изворима из области. Кандидат је потом анализирао најзначајније атрибуте за предвиђања модела и установио већ познате правилности познате из литературе што улива поверење да је креирани модел заиста успео да екстрахује релевантне информације из протеинских секвенци. Даље су разматрани потенцијални начини за даље унапређење модела, из чега се може закључити да кандидат одржава корак са достигнућима из области и већ прави планове за своја будућа надовезана истраживања. Дискусија се наставља концизним образлагањем поступка рада у развијеном софтверу за идентификацију и анализу HRGP секвенци, на које се надовезује разматрање анотације 62 биљна протеома коришћењем *ragp* софтвера.

У другом потпоглављу анализирана је величина и разноликост HRGP суперфамилије протеина *S. erythraea*. Највећи број идентификованих прототипских секвенци припада класичним арабиногалактанским протеинима са или без гликозил-фосфатидилинозитолног сидра. Поред прототипских секвенци, идентификоване су и AGP секвенце кичице које имају малу заступљеност карактеристичних мотива, а најфреквентнији домени у овим секвенцама су већ повезани са AGP на основу биоинформатичких или експерименталних доказа из литературе.

Треће потпоглавље дискусије се односи на изабраних 18 AGP гена чија је структура и експресија детаљније анализирана. Кандидат прво разјашњава избор управо ових гена, позивајући се на литературне изворе. Секвенце са фасциклинским доменом су одабране за анализу пошто су представници добро проучене фамилије протеина са претпостављеном улогом у адхезији ћелија, а за једану од изабраних секвенци је претходно показано да има пораст експресије током морфогенезе кичице гајене у условима *in vitro*. AGP са рецептор киназним доменом су одабране за испитивање експресије пошто је овај тип домена тек од скоро повезан са AGP карактеристичним мотивима, а према анотацији HRGP са *ragp* пакетом овај тип домена јако често поседује AGP карактеристичне мотиве у екстрацелуларном региону. Арабиногалактански пептиди су изабрани за детаљнију анализу пре свега пошто је њихова идентификација на основу хомологије проблематична па су у литератури најмање проучавани. Иако сама анализа експресије не може да укаже на

механистичку улогу производа гена, она може да укаже да ген на неки начин учествује у одређеном физиолошком процесу. У литератури је имплицирана улога AGP у одговору на повреде код различитих биљних врста, пре свега на основу промене експресије гена који их кодирају. Промена експресије два арабиногалактаснска пептида у одговору на повреде дискутована је кроз њихову потенцијалну улогу екстрацелуларних сензора. Даље је разматран утицај *Yariv* реагенса на експресију AGP гена. Утицај је испитиван пошто овај реагенс преципитира арабиногалактанске протене, а у литератури показано да он доводи инхибиције соматске ембриогенезе и органогенезе кичице, и да индукује одговор биљке најсличнији одговору на механичке повреде. Поменуто представља један од доказа да арабиногалактански протени учествују у овим процесима. Ипак, у експерименталном систему који је овде коришћен ни један од одабраних AGP гена није показао значајније промене експресије у одговору на β GlcY. На крају дискусије кандидат је разматрао експресију у узорцима ткива и органа *in vitro* гајених биљака, ткива из различитих фаза органогенезе и соматске ембриогенезе као и ткива из природе покушавајући да на основу шаблона експресије и литературних података дође до закључака о потенцијалним улогама испитиваних гена током развојних процеса, односно до гена који повезују механичку повреду са регенеративним ћелијским репрограмирањем.

У оквиру поглавља **ЗАКЉУЧЦИ** је прецизно и јасно изнет преглед најзначајнијих закључака изведених на основу добијених резултата који су у складу са постављеним циљевима докторске дисертације:

1. Проценом перформанси модела за превиђање вероватноће хидроксилације пролина на основу локалне секвенце биљних протеина, коришћењем четири алгорита машинског учења: KNN, SVM, RF и XGB у комбинацији са три типа одабира атрибута: mRMR, IGr и sfs, показано је да комбинација sfs и XGB алгоритама најбоље генерализује.
2. Тренирани модели машинског учења за превиђање вероватноће хидроксилације пролина представљају централни део нове методологије за идентификацију и анализу HRGP и AGP секвенци инкорпориране у R пакет *ragp* који је доступан за коришћење свима са основним знањем R програмског језика. Даље унапређење прецизности модела машинског учења за предвиђање позиција хидроксипролина у секвенцама биљних протеина ће неминовно унапредити специфичност и сензитивност идентификованих секвенци.
3. Развијена методологија је примењена за идентификацију HRGP и AGP секвенци у великом броју биљних протеома укључујући и протеом кичице. Идентификован је велики број и хетерогеност химерних AGP секвенци, што потенцијално значи да О-гликозилација на хидроксипролинима код биљака није ограничена на мањи скуп прототипских секвенци. Идентификовани HRGP гени представљају предуслов и ресурс за будућа истраживања улога чланова ове протеогликанске фамилије.
4. Испитивањем утицаја растућих концентрација β GlcY у хранљивим подлогама на експресију одабраних AGP гена, може се закључити да је додаток β GlcY индуковао релативно мале промене у експресији испитиваних гена, углавном без статистичке значајности, па на основу добијених резултата није могуће донети закључке о потенцијалној улози испитиваних гена.
5. Испитивањем експресије одабраних AGP гена након механичке повреде ткива показано је да је *CeAGp6*, који кодира 18 аминокиселина дуг AG пептид, јако

- индукован након повреде листа, а репримиран након повреде корена. Било би занимљиво функционално окарактерисати овај ген како би се утврдила његова улога током одговора биљног ткива на механичку повреду.
6. FLA транскрипти, *CeFLA1* и *CeFLA4* су најзаступљенији међу узорцима из природе, посебно у стаблу па је могуће да доприносе чврстости и еластичности стабла као што је претходно показано за њихове хомологе из *Arabidopsis thaliana*.
 7. Ниједан од изабраних 18 гена чија је експресија испитивана, није показао значајно повећање експресије у ткивима из различитих фаза органогенезе у условима *in vitro*. Према томе је вероватно да изабрани скуп од 18 *AGP* гена није укључио гене са значајном улогом током овог процеса.
 8. *CeFLA1*, који је показао повећање експресије током 48 h након повреде листа, а индукован је и у ембрионом калусу, могао би бити део мреже која повезује повреде и соматску ембриогенезу.

Поглавље ЛИТЕРАТУРА садржи 273 библиографске јединице из међународних и домаћих извора. Наведени извори покривају све области овог истраживања и на адекватан начин су наведени у тексту докторске дисертације.

Радови и конгресна саопштења из докторске дисертације

Б1. Радови у часописима међународног значаја

Dragičević, M., Paunović, D., Bogdanović, M., Todorović, S., Simonović, A. 2020. ragg: Pipeline for mining of plant hydroxyproline-rich glycoproteins with implementation in R. *Glycobiology* 30(1):19-35. 10.1093/glycob/cwz072. **M21, IF 4.194 (77/299 Biochemistry & Molecular Biology, 2018)**, <https://doi.org/10.1093/glycob/cwz072>

Paunović, D., Ćuković, K., Bogdanović, M., Todorović, S., Trifunović-Momčilov, M., Subotić, A., Simonović, A., Dragičević, M. The Arabinogalactan Protein Family of *Centaureum erythraea* Rafn. *Plants* 2021(10):1870. 10.3390/plants10091870. **M21, IF 3.935 (47/235 Plant Sciences, 2020)**, <https://pubmed.ncbi.nlm.nih.gov/34579403/>

Б2. Конгресна саопштења на скуповима међународног значаја

Paunović D., Bogdanović M., Trifunović Momčilov, M., Todorović, S., Simonović, A., Subotić A., Dragičević M. (2018) Are receptor tyrosine kinases chimeric AGPs?, 3rd International Conference on Plant biology, Book of Abstracts, Serbian Plant Physiology Society, pp. 17 - 17, isbn: 978-86-912591-4-3 (SPPS), Beograd, 9. - 12. Jun, 2018 **M34**

Paunović D., Bogdanović M., Todorović, S., Simonović, A., Dragičević, M. ragg: An R toolbox for mining Hydroxyproline rich glycoproteins, *Biologia Serbica: BELBI2018 Book of Abstracts*, Faculty of Sciences, University of Novi Sad, Serbia, vol. 40, no. 1, pp. 89 - 89, issn: 2334-6590, Beograd, 18. - 22. Jun, 2018 **M34**

Провера оригиналности докторске дисертације

На основу Правилника о поступку провере оригиналности докторских дисертација које се бране на Универзитету у Београду и налаза у извештају из програма iThenticate којим је извршена провера оригиналности докторске дисертације „Идентификација *AGP* гена кичице (*Centaurium erythraea*, *Gentianaceae*) и праћење њихове експресије у одговору на механичке повреде биљног ткива гајеног *in vitro*”, аутора Данијеле М. Пауновић, констатујем да утврђено подударане текста износи 9 %. Овај степен подударности последица је претходно публикованих резултата докторандових истраживања, који су проистекли из њене дисертације, цитата, математичких формула, делова у Материјалу и методама који су слични са одговарајућим деловима теза у којима се користи култура биљног ткива гајеног *in vitro* и молекуларно биолошке методе, и сличности теме истраживања са тезом др Биљане Филиповић под насловом „Улога арабиногалактанских протеина у процесу морфогенезе *in vitro* код кичице (*Centaurium erythraea* Rafn.)” на чије се резултате теза аутора Данијеле М. Пауновић ослања па је самим тим дат преглед, добрим делом, исте литературе у Уводу. Резултати из тезе др Биљане Филиповић су прописно цитирани. Треба додати да у одељцима Резултати и Дискусија практично нема подударности текста, осим са публикованим резултатима докторандових истраживања.

На основу свега изнетог, а у складу са чланом 8. став 2. Правилника о поступку провере оригиналности докторских дисертација које се бране на Универзитету у Београду, изјављујем да извештај указује на оригиналност докторске дисертације, те се прописани поступак припреме за њену одбрану може наставити.

Мишљење и предлог Комисије

На основу детаљног прегледа и анализе докторске дисертације Данијеле Пауновић под насловом „Идентификација *AGP* гена кичице (*Centaurium erythraea*, *Gentianaceae*) и праћење њихове експресије у одговору на механичке повреде биљног ткива гајеног *in vitro*”, која је реализована на одељењу за физиологију биљака Института за биолошка истраживања „Синиша Станковић“, Комисија сматра да је ова дисертација урађена у складу са свим принципима доброг научног рада и да представља оригинално научно дело са значајним научним доприносом областима физиологије, молекуларне биологије и биоинформатике биљака. Кандидаткиња Данијела Пауновић је остварила све циљеве наведене у пријави докторске дисертације. Током рада на својој дисертацији, кандидаткиња је показала висок степен самосталности, познавање научне литературе, способност да адекватно осмисли, планира и спроведе лабораторијска и *in silico* истраживања, а да резултате критички анализира и изводи правилне закључке. Посебан значај овој дисертацији даје савремен приступ који је уврстио методе машинског учења и креирање нових информатичких алата који ће омогућити напредак у проучавању функција протеина ћелијског зида биљака.

Комисија са задовољством предлаже Наставно-научном већу Биолошког факултета Универзитета и Београду да прихвати позитиван извештај о урађеној докторској дисертацији Данијеле Пауновић под насловом „Идентификација *AGP* гена кичице (*Centaurium erythraea*, *Gentianaceae*) и праћење њихове експресије у одговору на механичке повреде биљног ткива гајеног *in vitro*” и тиме кандидаткињи омогући јавну одбрану докторске дисертације.

КОМИСИЈА:

У Београду, 21.04.2022.године

др Тијана Цветић Антић, ванредни професор,
Универзитет у Београду, Биолошки факултет

др Марко Ђорђевић, ванредни професор,
Универзитет у Београду, Биолошки факултет

др Марија Марковић, виши научни сарадник,
Институт за биолошка истраживања
„Синиша Станковић“ -Институт од
националног значаја за Републику Србију