

Pētniecības pieteikuma vienošanās Nr. *1.1.1.2/VIAA/1/16/094*

Pusgada populārzinātnisks pārskats (1.03.2018.-31.08.2018.) par projekta īstenošanas gaitu “Transponējamo elementu variāciju izpēte parastās priedes (*Pinus sylvestris* L.) gēnu rajonos”

Turpinājās priežu referenču genomu bioinformātiskā analīze pēc iepriekš izstrādātās darba plūsmas, un proti, veikta gēnu flankējošo sekvenču analīze *P.taeda* v.1.0. un *P.lambertiana* genomiem. Izveidota apkopojošā tabula ar flankējošo rajonu kopīgo skaitu un atrasto LTR līdzību attiecību, kas parādīja, ka pirmās versijas genomiem mobilo elementu sadalījums flankējošās gēnu sekvencēs atšķiras no *P.taeda* v.2.0. genoma. Atšķirībā no otrās versijas genoma rezultātiem, kur mobilo elementu sadalījums ir vienmērīgs visos rajonos, kas atrodas tālāk par 2 kB no gēniem, pirmajās versijās mobilo elementu skaits graduāli samazinās katrā rajonā līdz ar attāluma samazināšanos līdz gēnam. Turklāt, pirmajām genoma versijām gēnu kopas ir sadalītas augstās kvalitātes un zemās kvalitātes gēnos, un augstās kvalitātes gēniem mobilo elementu skaits flankējošās gēna daļās ir niecīgs. Šie rezultāti un zināšanas par citu augu genomu struktūru, norāda uz to, ka mobilie elementi ar augstu iespējamību tika izdzēsti no augstās kvalitātes gēnu sekvencēm. Šāda prakse ir normāla NGS (jaunās paaudzes sekvenēšanas tehnoloģijas) veidotu genomu salikšanā. Mobilie elementi atkārtojas priežu genomā tūkstošiem reižu un to sekvenču ir ļoti līdzīgas savā starpā, tāpēc šo atkārtojumu precīzās vietas noteikšana bieži vien nav iespējama. Otrās versijas genoma salikšanā tiek izmantotas garāki sekvenēšanas produkti (vidēji 5 kB), kas ļauj zinātniekiem noteikt vietu mobila elementa daļai, taču elementa pilna izmēra vai himēriskā elementa (tāda, kas sastāv no dažādu elementu daļām), paliek nezināma, jo mobilo elementu garums vidēji sasniedz 15 kB un vairāk. Tas izskaidro arī to, kādēļ *P.taeda* v.2.01. genomam gēnu nekodējošās daļas ir daudzas reizes garākas, kā tas bija genoma pirmajā versijā. Projekta plānošanas laikā publiski bija pieejama tikai mazāk precīza pirmā genoma versija, datubāzes joprojām satur gēnu īsās versijas, kas ļautu bez grūtībām analizēt šādas sekvenču. Tādēļ otrās genoma versijas analīzes rezultātā var secināt, ka turpmāk ir iespējams saskarties ar sarežģījumiem un plānu izmaiņām, jo ļoti garas gēnu intronu sekvenču nebūs iespējams analizēt ar iepiņanoto metožu kopu.

Lai turpmāk saprast atrasto gēnu kopu nozīmi, radās nepieciešamība pēc genoma funkcionālas anotācijas. Gēnu ontoloģijas (saīsināti GO) analīze ir gēnu sekvenču veidotā proteīna vai RNS produkta apraksts, kas var ietvert informāciju par molekulas zināmu dalību bioloģiskā procesā, molekulārās funkcijas izpildē, kā arī atrašanās vietu. Daži gēni satur aprakstu visās trijās grupās, citiem ir pieejama informācija par dalību vienā no grupām, daudzu citu augu gēnu produktu funkcijas

joprojām nav zināmas. Katrā grupā ir vairākas apakšgrupas ar unikālu numuru (GO-kategorija), kas raksturo gēna produkta dalību noteiktajā procesā ar noteiktu precizitātes līmeni no vispārīga apraksta (piem. dalība bioloģiskajā regulācijā vai proteīnu sintēzē), līdz vispārīgai gēna produkta funkcijai (lignīna biosintēze, Kalcija jonu transports utml). Izmantojot šos numurus ir iespējams sašķirot gēnus pēc to produktu iesaistes procesos un saprast, vai viņiem varētu būt kopīga darbība. Lai piešķirtu kādu no funkcionālām GO-kategorijām, tiek izmantoti gēnu transkripti, jo tie nesatur nekodējošās gēna daļas. Priedes genoma pirmajām versijām ir pieejama gēnu funkcionāla analīze, kas tika balstīta uz izpētīto augu gēnu zināšanām. Tomēr šī anotācija mūsu gadījumā nevarētu tikt izmantota bez iepriekšējās apstrādes, jo gēnu numerācija abām versijām atšķiras. Lai izveidot jaunu gēnu anotāciju ir pieejamas maksas programmas, kas darbojās ļoti ilgi, bet rezultāti nav izmantojami, jo arī priežu genoma transkripti tomēr var saturēt mobilu elementu daļas, kas liek programmai automātiski piešķirt visam gēnam citu funkciju. Lai varētu šīs kļūdas novērst, izmantoju pieejamu precīzāku gēnu anotāciju un ar dažādām manipulācijām izveidoju *P.taeda* v.2.01. un *P.lambertiana* v.1.0. genomu pilnīgu genoma anotāciju. Rezultātus varētu iespaidot arī tas, ka pusei no visu priežu gēniem nav pieejama anotācija, jo šie gēni neuzrāda līdzību kādam zināmam produktam. Taču, iespējams, tieši šie gēni ir iesaistīti procesos, kas nosaka priedes sugu specifiskās pielāgošanās spējas. Turpmāk tika izstrādāta darba plūsma gēnu kopu analīzei ar *Cytoscape* programmas *BINGO* lietotni. *BINGO* lietotne ļauj pašizgatavoto datu ievadi, kas ļauj izmantot visjaunāko funkcionālo gēnu anotāciju. Tāpat, šī lietotne paaugstina analīzes precizitāti, jo ir iespējams izmantot visa genoma gēnu grupu anotēšanas analīzi. Tomēr datu formāts neļāva iedarbināt lietotni uzreiz un nācās sazināties ar lietotnes veidotāju. Rezultātā ieguvu iespēju analizēt iepriekš atrasto gēnu kopu funkcionālu saistību ar gēnu tīklu palīdzību. Turklāt, funkcionālo gēnu grupu reprezentācija kopā tiek statistiski novērtēta pēc esošās visa genoma gēnu grupu anotācijas, kas ir ļoti svarīgi nepilnīgās anotācijas gadījumā. Tika atrasti saistīti gēnu tīkli, kas piedalās auga aizsardzības mehānismu noteikšanā dažādos apstākļos, šūnu līdzsvara uzturēšanā, šūnapvalka sintēzē, hormonu sintēzē, proteīnu transporta nodrošināšanā u.c. Šīs funkcijas tika apkopotas tabulā, kas ļauj izvērtēt katra gēnu tīkla iespējamu atšķirību un nozīmi.

Mobilie elementi var saturēt dažādus signālus, kas spēj piesaistīt molekulas, kas veidojas atbildot uz apkārtējās vides izmaiņām, un tādā veidā regulēt gēnu darbību. No literatūras ir zināms, ka piemēram, metilēšanas signāli gēna nekodējošās sekvencēs spēj inaktivēt visa gēna sekvenci. Ja sekvence satur papildus signālus, kas aktivē gēna darbību, tad šādas sekvences var pozitīvi ietekmēt produkta daudzumu noteiktajos apstākļos. Tādā veidā, transponējamie elementi (TE), izplatot šādus signālus gēnu rajonos, varētu palīdzēt augam ātri pārslēgt savu metabolismu no ierastās darbības uz ārkārtas situācijas režīmu. Tas varētu izskaidrot šo elementu daudzveidību, jo dažādos apstākļos vai patogēnu ietekmē laika gaitā

var rasties nepieciešamība pēc dažādiem aizsardzības mehānismiem; TE skaitu un nevienmērīgo izplatību gēnu rajonos, jo tikai noderīgi pārkārtojumi tiks saglabāti dabiskās izlases procesā; TE korelāciju ar sugas izcelsmi, jo katra suga ir pielāgojusies noteiktiem augšanas apstākļiem; auga spēju dinamiski mainīties, ieslēdzot vai izslēdzot nepieciešamos gēnus no atbildes tīkla, izmainot mazāk vērtīgas nekodējošās sekvences. Tādēļ, identificētās būtiski palielinātās TE sekvences tika analizētas izmantojot bioinformātiskus rīkus augu transkripcijas faktoru piesaistes vietu noteikšanai un citu funkcionālo motīvu analīzei, līdzības noteikšanai ar citiem mobiliem elementiem. Dati tika apkopoti tabulā. Identificēju interesantu signālu-(AGNN)₃(AG)₃(NNAG)₂- kurš ir ļoti līdzīgs labi izpētītam citu augu gaismas-responsīvam elementam, tomēr priekšu mobilais elements satur garāku atkārtojumu.

Tāpat, otrajā pusgadā ir uzsākts darbs pie publikācijas, tika apkopoti dati par atrastajiem mobiliem elementiem, izveidoti attēli. Publikācijas melnraksts tika nosūtīts zinātniskajam darba konsultantam Dr.Biol. Dainim Ruņģim. Turpinās darbs arī pie datu pārbaudes, lai izslēgtu iespējamās kļūdas publicētos datos. Ir jāpārbauda atrasto mobilo elementu sekvences, vai tie nepieder viena atkārtojuma dažādām daļām.

Šī pusgada laikā arī piedalījos zinātnes popularizēšanas un tīklošanās pasākumos. Sagatavots stenda referāts “*Distribution of predicted Retrotransposon LTRs in gene flanking regions of the Pinus taeda genome v.2.0.*” (līdzautori: Angelika Voronova, Martha Rendon, Par Ingvarsson, Dainis Ruņģis). No 4/07/2018 līdz 8/07/2018 esmu piedalījies starptautiskajā konferencē “Evolūcija. Ģenētiskie jaunumi/ Genoma variācija caur RNS tīkliem un vīrusiem” (*Evolution. Genetic Novelty/ Genomic Variations by RNA Networks and Viruses*), kas notika Salzburgā (Austrija). Komandējuma atskaite un informācija par dalību tika publicēta www.silava.lv mana projekta sadaļā. Sagatavota populārzinātniskā prezentācija “*Non-coding genome dynamics in stress conditions*”, ar kuru uzstājos Latvijas un Lietuvas Jauno zinātnieku apvienību rīkotajā pasākumā “*International smithy of ideas 2018*”, kas notika Lietuvā 24.-26.08.2018.

Pārskatu sagatavoja: Angelika Voronova