

Gēnus regulējošo sekvenču izpēte priedes genomā

LVMI Silava pētniece Angelika Voronova 2017. gada septembrī uzsāka pēcdoktorantūras pētījumu “Transponējamo elementu variāciju izpēte parastās priedes (*Pinus sylvestris* L.) gēnu rajonos”. Par pētījumā iegūtajiem pirmajiem rezultātiem un turpmāk plānoto stāsta ANGELIKA VORONOVA.



ŠIS pētījums ir turpinājums iepriekšējam darbam par transponējamiem elementiem (TE) priedes genomā. Iepriekš izzinājām šo elementu aktivēšanos, atbildot uz dažādiem stresu izraisošiem faktoriem, to izplatību priedes ģints sugās, kā arī citos kailsēkļos. Paaugstināta šo sekvenču frekvence ir novērota dažādos līmeņos: gan vietējai priežu populācijai, kas aug atšķirīgos ūdens režīma apstākļos, gan priežu sugām, kuras ir adaptējušās ekstrēmiem vides apstākļiem.

TE ir sekvences, kas kādreiz spējušas pārvietoties genomā vai pārvietoja pašlaik; pārvietošanās biežāk notiek pārmērīgas audos vai arī atbildot uz stresa faktoriem. Proporcioniāli TE aizņem daudz lielāku genoma daļu nekā gēni, un šīs sekvences ir ļoti daudzveidīgas. Priedes genoms var saturēt vairāk nekā 100 tūkstošu viena tipa atkārtojumu, kas ir izplatīti visās hromosomās. Spēju aktivēt stresu apstākļos tiem piešķir specifiskie signāli vai DNS sekvenču motīvi. Pašlaik zinātnieki uzskata, ka TE ir nozīmīgs rīks jauno ģenētisko izmaiņu veidošanai nelabvēlīgos vai mainīgos apstākļos.

Pētījuma pirmais posms bija veltīts bioinformātikai priežu references genomu (*Pinus taeda* un *Pinus lambertiana*) analīzei. Lai operētu ar apjomīgo genoma datu daudzumu, izmantoju augstas veiktspējas skaitļošanas klasteri, kas bija pieejams Zviedrijā. Upsalas Biocentrā mācījos pielietot dažādus skaitļošanas algoritmus, pielāgojot tos sava pētījuma mērķiem. Izolēju 36 730 gēnu sekvences un analizēju, kuru ģeni satur TE sekvences, to lokalizāciju, TE tipus un tml. Salīdzināju iegūtos datus starp dažādiem genomiem un genomu versijām. Zviedrijā bija pieejama arī *P. sylvestris* nepublicēta genoma sākuma versija, taču analīzes parādīja, ka šīs sekvences nesatur pilna izmēra gēnus, kā arī atkārtojumus.

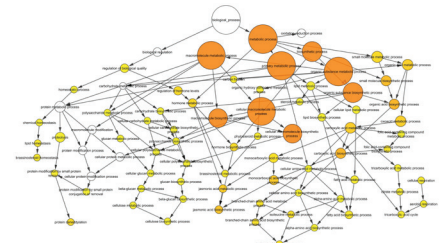
Atgriezoties Latvijā, datus nācās vairākkārtīgi pārbaudīt un atkārtoti analizēt, kā arī rast jaunas datu analizēšanas pieejas. Šim nolūkam sākām sadarboties ar Rīgas Tehniskās universitātes HPC skaitļošanas centru, kur iegūtā pieredze ļāva veikt visu specifisko programmu instalēšanu un sekmīgi

darboties. Piemēram, ideju par TE sekvenču izplatības analīzi, atkarībā no attāluma līdz gēna sekvencai, realizēju jau patstāvīgi. Atradu TE sekvences, kas biežāk ir sastopamas gēnam tuvākos rajonos (p-vērtība <0,0001), kas liecina par selektīvā spiediena darbību uz šīm sekvencēm. Vēl viens izaicinājums bija izveidot un pielāgot visu gēnu funkcionālo anotāciju, lai varētu veikt gēnu tīklu analīzi (skat. att.). Funkcionāla atrasto gēnu kopu analīze ļauj saprast, kādos procesos ir iesaistīti gēnu veidotie produkti un vai kāds no TE varētu veidot efektīvu saistību. Tāpat tika analizēti atrasto TE DNS motīvi, meklējot augiem nozīmīgus signālus, kas varētu ierosināt vai bloķēt blakus esošo gēnu darbību. Atrasti ļoti interesanti motīvi, kuru piedalīšanās pierādīta stresa atbilstu regulācijā citos augos. Iegūto datu apjoms ir ļoti liels, bet ceru drīzumā visus apkopot. Viena auga genoma sekvence neatklāj visu populācijā esošo ģenētisko daudzveidību. Tāpat, genoma dati joprojām neļauj precīzi noteikt pilna izmēra vai daļēju TE lokalizāciju. Tādēļ pētījuma turpinājumā no datora monitora būs jāatgriežas pie darba laboratorijā.



Parastās priedes sējeņu inokulācijas eksperiments ar *H. annosum*, kas tika izmantots TE ekspresijas pētījumam.

Izvēlēšos dažus no atklātiem gēnu tīkliem, lai praktiski pārbaudītu, vai atklātās izmaiņas populācijā ir stabilas, polimorfismu variāciju, vai saistības ir atrodamas arī parastās priedes gēniem.



Gēnu saistību tīkls, kuru varētu ietekmēt identisks TE (attāluma koordinātes no 0 līdz 1000 bp).