

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛАРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра молекулярной биологии**

Аннотация к дипломной работе

**НЕНАХОВА
ТАТЬЯНА ВАДИМОВНА**

**МОЖЕТ ЛИ ЭКСПРЕССИОННАЯ МОДЕЛЬ ПРЕДСКАЗАТЬ
КАЧЕСТВО ВОЛОКНА ЛЬНА?**

Научный руководитель:
кандидат биологических наук
доцент Ю.В. Полухович

Минск 2022

РЕФЕРАТ

Дипломная работа: 47стр., 3 рис., 13 табл., 24 источника.

Ключевые слова: экспрессионная модель, лён-долгунец, селекция, генетическое предсказание.

Объект исследования: экспрессионные модели предсказывающие параметры качества льна-долгунца LASSO и Stepwise.

Цель: проверка одного из подходов, который позволяет установить связь между уровнем экспрессии генов, рассматриваемых в качестве потенциальных маркеров, и тремя признаками, характеризующими качество, одной из сельскохозяйственных культур – льна-долгунца.

Методы исследования: молекулярно-генетические (выделение РНК, полимеразная цепная реакция), биоинформатические.

В результате проведённого исследования были проанализированы параметры качества одного сортообразца и четырёх сортов льна-долгунца. Для этих образцов на основании экспрессионных данных, полученных в результате полимеразной цепной реакции, были рассчитаны прочность на разрыв, гибкость и техническая длина стебля, при помощи двух экспрессионных моделей Stepwise и LASSO. Далее проводилась проверка на аккуратность предсказания этих моделей, путём сравнения эмпирических и экспериментальных данных.

Данные полученные методом LASSO, наиболее приближены к эмпирическим измерениям. Данная модель имеет наилучшие значения коэффициентов корреляции, по 2 из 3 исследуемых параметров: гибкость и техническая длина стебля. При этом она соблюдает некоторое ранжирование по всем параметрам. Этот метод регуляризированной регрессии имеет потенциал для дальнейшего развития и модификации. Хуже показала себя модель Stepwise. Несмотря на то, что она хорошо предсказала независимую переменную характеризующую прочность на разрыв, а также наилучший коэффициент корреляции поэтому же параметру, остальные показатели были слишком сильно занижены и при этом наблюдалась сильная девиация параметров.

Работа направлена на определение наилучшей экспрессионной модели для предсказания параметров качества льняного волокна при дальнейшем использовании в селекции, с целью получения наибольшего экономического эффекта в льноводческой отрасли сельского хозяйства.

SUMMARY

Diploma project: 47 pages, 3 figures, 13 tables, 24 sources.

Key words: expression model, flax, selection, genetic prediction.

The object of research: expression models predicting the quality parameters of flax LASSO and Stepwise.

The aim of the research: to test one of the approaches that makes it possible to establish a link between the level of expression of genes considered as potential markers and three characteristics characterizing the quality of one of the agricultural crops – flax.

The research methods: molecular genetic (RNA isolation, polymerase chain reaction), bioinformatic.

As a result of the study, the quality parameters of one variety and four varieties of flax were analyzed. For these samples, based on the expression data obtained as a result of polymerase chain reaction, the tensile strength, flexibility and technical length of the stem were calculated using two expression models Stepwise and LASSO. Next, the accuracy of the prediction of these models was checked by comparing empirical and experimental data.

The data obtained by the LASSO method are the closest to empirical measurements. This model has the best values of correlation coefficients, according to 2 of the 3 parameters studied: flexibility and technical length of the stem. At the same time, it observes some ranking in all parameters. This method of regularized regression has the potential for further development and modification. The Stepwise model proved to be worse. Despite the fact that it predicted well the independent variable characterizing the tensile strength, as well as the best correlation coefficient for the same parameter, the remaining indicators were too much underestimated and at the same time there was a strong deviation of the parameters.

The work is aimed at determining the best expression model for predicting the quality parameters of flax fiber for further use in breeding, in order to obtain the greatest economic effect in the flax-growing branch of agriculture.

РЭФЕРАТ

Дыпломная работа: 47стр., 3 мал., 13 табл., 24 крыніцы.

Ключавыя словы: экспрэсійная мадэль, лён-даўгунец, селекцыя, генетычнае прадказанне.

Аб'ект даследавання: экспрэсійныя мадэлі прадказваюць параметры якасці лёну-даўгунца LASSO і Stepwise.

Мэта: праверка аднаго з падыходаў, які дазваляе ўсталяваць сувязь паміж узроўнем экспрэсіі генаў, якія разглядаюцца ў якасці патэнцыйных маркераў, і трыма прыкметамі, якія характарызуюць якасць, адной з сельскагаспадарчых культур-лёну-даўгунцу.

Метады даследавання: малекулярна-генетычныя (вылучэнне РНК, палімеразнай ланцуговая рэакцыя), біяінфарматычныя.

У выніку праведзенага даследавання былі прааналізаваныя параметры якасці аднаго сортообразца і чатырох гатункаў лёну-даўгунцу. Для гэтых узораў на падставе экспрэсійных дадзеных, атрыманых у выніку палімеразнай ланцуговай рэакцыі, былі разлічаны трываласць на разрыў, гнуткасць і тэхнічная даўжыня сцябла, пры дапамозе двух экспрэсійных мадэляў Stepwise і LASSO. Далей праводзілася праверка на акуратнасць прадказання гэтых мадэляў, шляхам параўнання эмпірычных і эксперыментальных дадзеных.

Дадзеныя атрыманыя метадам LASSO, найбольш набліжаны да эмпірычным вымярэннях. Дадзеная мадэль мае найлепшыя значэння каэфіцыентаў карэляцыі, па 2 з 3 доследных параметраў: гнуткасць і тэхнічная даўжыня сцябла. Пры гэтым яна выконвае некаторы ранжыраванне па ўсіх параметрах. Гэты метады рэгулярызаванай рэгрэсіі мае патэнцыял для далейшага развіцця і мадыфікацыі. Горш паказала сябе мадэль Stepwise. Нягледзячы на тое, што яна добра прадказала незалежную зменную якая характарызуе трываласць на разрыў, а таксама найлепшы каэфіцыент карэляцыі таму ж параметры, астатнія паказчыкі былі занадта моцна прыніжаныя і пры гэтым назіралася моцная дэвіяцыя параметраў.

Праца накіравана на вызначэнне найлепшай экспрэсійнай мадэлі для прадказання параметраў якасці льнянога валакна пры далейшым выкарыстанні ў селекцыі, з мэтай атрымання найбольшага эканамічнага эфекту ў льнаводчай галіны сельскай гаспадаркі.