



УДК 577.21:633.52

ПОШУК І АНАЛІЗ ПОСЛІДОВНОСТЕЙ ГЕНІВ

АКТИНУ В ГЕНОМІ ЛЬОНУ

А. С. ПОСТОВОЙТОВА, аспірант*

Г. Я. БАЄР, кандидат біологічних наук

М. О. ПИДЮРА, інженер

Н. Л. ПАСТУХОВА, кандидат біологічних наук

Я. В. ПІРКО, кандидат біологічних наук

А. І. ЄМЕЦЬ, доктор біологічних наук

Я. Б. БЛЮМ, доктор біологічних наук

ДУ «Інститут харчової біотехнології та геноміки НАН України»

E-mail: nastya.postovoytova@gmail.com

***Анотація.** Здійснено біоінформаційний пошук гомологів генів актину в геномі льону і проведений аналіз відібраних послідовностей. З десяти гомологів генів, що кодують актин більшість має схожу екзон-інтронну структуру – переважно чотири екзони та три інтрони. Встановлено ступінь ідентичності повних послідовностей генів, їх кодуючих ділянок та амінокислотних послідовностей, який становить в середньому – відповідно 80 %, 90,5 % та 97 %. Отримані дані свідчать, що екзони належать до найбільш консервативних ділянок генів, а послідовності інтронів мають суттєві відмінності, тобто їм притаманний високий ступінь поліморфізму.*

***Ключові слова:** льон, актин, локус*

Льон (*Linum usitatissimum* L.) є однією з основних технічних культур, яка здавна забезпечує потреби людини у високоякісному волокні. [1]. Основним полімером льоноволокна є целюлоза, вміст якої може досягати 70 % від маси зрілого волокна. Цей полімер переважно накопичується в клітинах флоєми (луб'яне волокно), де відкладається в клітинній стінці [2]. Припускають, що

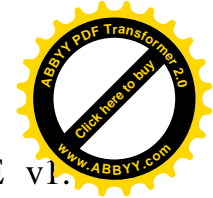
*Науковий керівник – кандидат біологічних наук, старший науковий співробітник Я. В. Пірко



мікротрубочки та актинові філаменти слугують не лише «матрицею» для Р₂ целюлосинтазних комплексів [3, 4, 5, 6], а є і орієнтиром для їх вбудовування та видалення із плазматичної мембрани [5, 7], що створює передумови для диференціації сортів льону за їх агротехнічними властивостями, а саме здатністю утворювати льоноволокно. Раніше проводилися роботи щодо пошуку й аналізу генів целюлосинтазного комплексу [8] і генів актину, α – та β -тубуліну в геномах вищих видів рослин та льону, зокрема [9, 10]. Подальше вивчення структури генів білків цитоскелету, зокрема генів, що кодують синтез актину, дослідження механізмів їх коекспресії відкриває широке коло можливостей для вивчення генетичної диференціації сучасних сортів льону.

Мета дослідження – проведення біоінформаційного пошуку й аналізу генів, що кодують актин (основний білок мікрофіламентів), в геномі льону для подальшого застосування інформації про структуру генів у селекційно-генетичних дослідженнях.

Матеріали та методика досліджень. Пошук генів, що кодують актин, здійснювали на основі анотованої послідовності актину льону (Q5I3E6_LINUS) за допомогою інструменту TBLASTN версії 2.2.26+ в базі даних Phytozome v9.1 (www.phytozome.net). Така методика використовувалась у попередніх роботах [8, 9] і передбачає пошук на основі порівняння заданих амінокислотних послідовностей із трансльованою у амінокислотні послідовності базою даних нуклеотидів. Для цього використали налаштування програми TBLASTN за замовчуваннями. Відібрали амінокислотні послідовності, для яких поріг відсіву випадкових співпадінь (e-value) складав не більше 1^{-10} . Отримані послідовності перевіряли за допомогою алгоритму Hidden Markov Model у програмі HMMER3 (www.hmmer.janelia.org) із використанням референтної бази даних родин білків Pfam-A (версія 25.0) [11]. Генетичні локуси, транскрибовані послідовності яких були віднесені за даними бази даних Pfam до родини актинів, визначили як гени актину льону. Для вирівнювання генетичних послідовностей і побудови філогенетичного дерева використовували програму Clustal W [12]. Ступінь



гомології (ідентичності) визначали за допомогою програми UGENE v1. [13].

Результати досліджень та їх обговорення. В базі даних Uniprot (www.uniprot.org) на даний момент знаходиться дві анотовані послідовності, закодовані у геномі льону (*Linum usitatissimum*), довжиною 375 і 208 амінокислотних залишків, які віднесені до фрагментів послідовностей актину. Середня кількість амінокислот у послідовностях актину *A. thaliana* складає близько 377 амінокислот [14]

В геномі льону було виявлено 10 генів актину. Вони розташовані в локусах Lus10006783, Lus10006784, Lus10001693, Lus10001694, Lus10005457, Lus10004956, Lus10016259, Lus10004169, Lus10040826 і Lus10005163.

Результати проведеного аналізу повних генних послідовностей свідчать про те, що гени актину в геномі *L. usitatissimum* містять переважно по три інтрони і чотири екзони (табл., рис. 1) окрім генів, закодованих у локусах Lus10004169 та Lus10001693, які мають по два інтрони і три екзони. Екзони в відібраних генах містять однакову кількість пар нуклеотидів: 60 (перший), 394 (другий), 614 (третій) та 66 (четвертий) відповідно. Виключенням є локуси Lus10001693 та Lus10004169. В структурі інтронів жодної системності не спостерігається. Загалом, як видно з таблиці, довжини інтронів складають від 66 до 481 п. н., а довжини екзонів – 60 - 761 п. н.

Екзон-інтронна структура генів актину *L. usitatissimum* (п. н.)

Локус	Перший екзон	Перший інтрон	Другий екзон	Другий інтрон	Третій екзон	Третій інтрон	Четвертий екзон	Загальна довжина гену
Lus10005163	60	346	394	87	614	94	66	1661
Lus10006783	60	78	394	79	614	481	66	1772
Lus10006784	60	89	427	66	614	105	66	1427
Lus10001693	319	86	614	99	66	-	-	1184
Lus10005457	60	87	394	300	614	93	66	1614
Lus10004956	60	106	394	355	614	85	66	1680
Lus10004169	60	88	394	295	761	-	-	1598
Lus10016259	60	91	394	321	614	89	66	1635
Lus10001694	60	321	394	87	614	94	66	1636
Lus10040826	60	95	394	114	614	79	66	1422

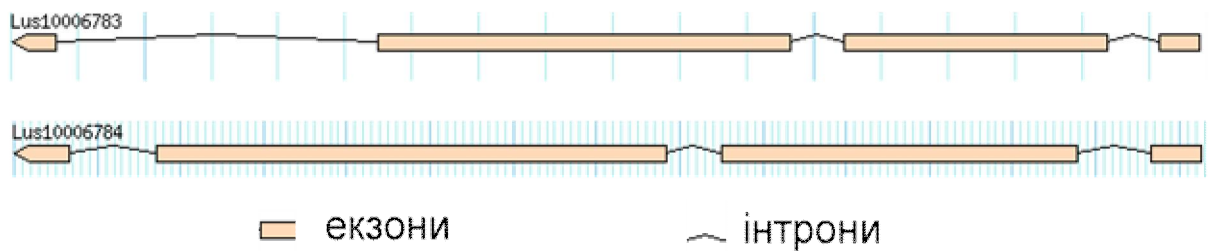


Рис. 1. Екзон-інтронна структура генів актину в локусах Lus10006783 і Lus10006784 (схематичне зображення)

Середня кількість амінокислот у послідовностях, що транлюються на підставі даних про гени актину *A. thaliana*, складає близько 377 амінокислот. Так, наприклад, гени Lus10006783 і Lus10006784 кодують майже ідентичні білкові продукти. Єдиною різницею є те, що у локусі Lus10006783 наявна одна унікальна вставка всередині послідовності, що складає 10 амінокислот. Але відрізняються ці локуси за довжиною ДНК-ланцюга на 345 нуклеотидів (довжина генів 1772 і 1427 нуклеотидів відповідно). Обидва гени містять чотири екзони і три інтрони, але довжина третього інтрона у гені Lus10006783 значно більша.

За результатами проведеного множинного вирівнювання та здійсненого аналізу показано, що ступінь ідентичності повних гомологічних послідовностей генів актину льону складає в середньому 80 %, кодуючих ділянок (лише екзони) - близько 90,5 % та амінокислотних послідовностей – 97 %. На основі подібності послідовностей було побудовано філогенетичне дерево (рис. 2).

З рис.2 видно, що послідовності в локусах Lus10006783 і Lus10006784, а також Lus10004956, Lus10005457 і Lus10040826 кодують практично ідентичні продукти, але Lus10004956 і Lus10006783 мають по одній унікальній вставці всередині послідовності. Амінокислотна послідовність актину, що транлюється з гомолога гена, закодованого в локусі Lus10004169, найбільше відрізняється від інших як за складом амінокислот, так і за їх кількістю. Це



безпосередньо пов'язано з екзон-інтронною структурою даного гену, яка значно відрізняється від структур інших послідовностей, а саме більшою кількістю пар нуклеотидів першого і другого екзонів та відсутністю четвертого (табл.).

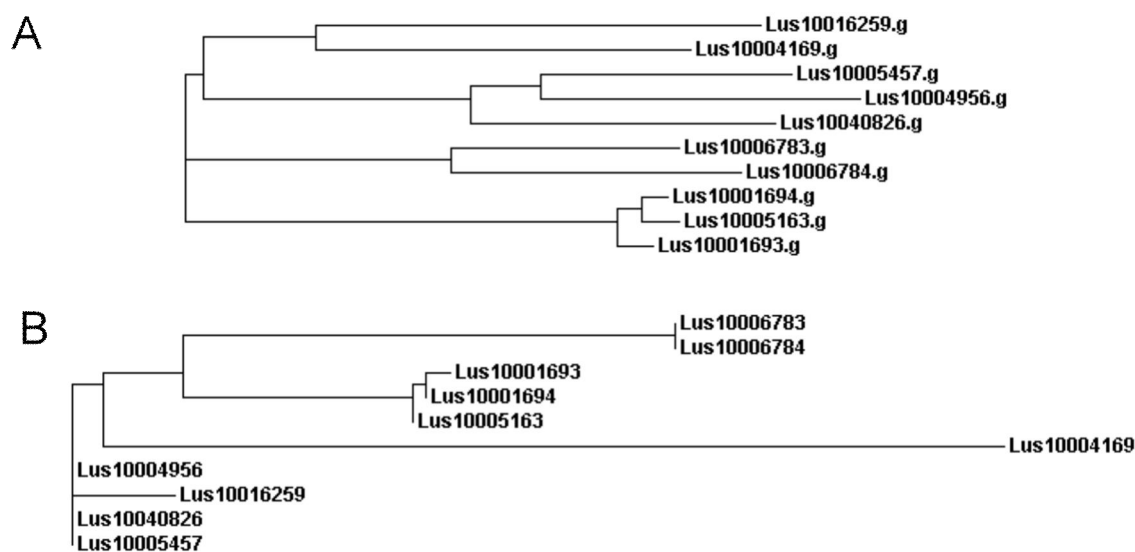
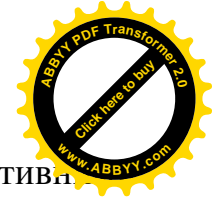


Рис. 2. А – філогенетичне дерево подібності визначених гомологів генів, що кодують різні форми актину льону; В – філогенетичне дерево подібності продуктів (АК послідовностей) гомологів генів, що кодують різні форми актину льону

Таким чином, гомологи генів різних форм актину льону, не зважаючи на високий ступінь ідентичності закодованих продуктів, мають значні відмінності за складом та довжиною інтронних фрагментів. Можливо, що в експресії генів різних форм актину льону задіяні складні регуляційні процеси, до яких можуть бути залучені інтрони. Щоправда, дане питання ще потребує додаткового дослідження.

Висновки

1. Визначено і проаналізовано 10 генів, що кодують актин в геномі льону.
2. Встановлено ступінь ідентичності повних послідовностей генів актину льону, кодуючих ділянок та амінокислотних послідовності на рівні відповідно 80 %, 90,5 % та 97 %.



3. Підтверджено, що саме екзони генів є найбільш консервативні ділянками, в той же час гени відрізняються значною мірою за рахунок інтронних фрагментів.

СПИСОК ЛІТЕРАТУРИ

1. Reddy N. Biofibers from agricultural byproducts for industrial applications / N. Reddy, Y. Yang // Trends in Biotech. – 2005. – 23. – P. 22-27.

2. Titok V. Leontiev V., Shostak L., Khotyleva L. Thermogravimetric analysis of the flax bast fiber bundle / V. Titok, V. Leontiev, L. Shostak, L. Khotyleva // J. Nat. Fibers. – 2006. – 3. – P. 35-41.

3. Paredez A. R. Visualization of cellulose synthase demonstrates functional association with microtubules / A. R. Paredez, C. R. Somerville, D. W. Ehrhardt // Sci. – 2006. – 312. – P.1491-1495.

4. Emons A. M. C. Microtubules and cellulose microfibrils: how intimate is their relationship? / A. M. C. Emons, H. Hofte, B. M. Mulder // Trends Plant Sci. – 2008. – 12. – P. 279-281.

5. Crowell E. F. Pausing of Golgi bodies on microtubules regulates secretion of cellulose synthase complexes in *Arabidopsis* / E. F. Crowell, V. Bischoff, T. Desprez, A. Rolland, Y. D. Stierhof, K. Schumacher, M. Gonneau, H. Hofte // Plant Cell. – 2009. – 21. – P.1141-1154.

6. Chan J. The rotation of cellulose synthase trajectories is microtubule-dependent and influences the texture of epidermal cell walls in *Arabidopsis* hypocotyls / J. Chan, E. Crowell, M. Eder // J. Cell Sci. – 2010. – 123. – P. 3490-3495.

7. Gutierrez R. Arabidopsis cortical microtubules position cellulose synthase delivery to the plasma membrane and interact with cellulose synthase trafficking compartments / R. Gutierrez, J. J. Lindeboom, A. R. Paredez, A. M. Emons, D. W. Ehrhardt // Nat. Cell Biol. – 2009. – 11. – P. 797-806.



8. Пыдюра Н. А. Биоинформационный поиск генов целлюлозосинтазы (*Linum usitatissimum*) и их филогенетический анализ / [Н. А. Пыдюра, Г. Я. Баер, Д. В. Галиновский и др.] // Цитология и генетика - 2015. - 49(5). – P. 3-12.
9. Рабокoнь А. Н Анализ гомологов генов основных белков цитоскелета у различных видов высших растений / [А. Н. Рабокoнь, А. С. Постовойтова, Я. В. Пирко, Я. Б. Блюм] // Фактори експериментальної еволюції організмів – 2014. – 14 – P. 71-76.
10. Баер Г. Я. Біоінформаційний пошук послідовностей генів, що кодують тубуліни у геномі льону / [Г. Я. Баер, М. О. Пидюра, Я. В. Пирко та ін.] // Фактори експериментальної еволюції організмів – 2014. – 14 – P. 14-17.
11. Punta M. The Pfam protein families database / M. Punta, P. C. L. Coggil, R. Y. Eberhardt, J. Mistry, J. Tate, C. Boursnell, N. Pang, K. Forslund, G. Ceric, J. Clements, A. Heger, L. Holm, E. L. L. Sonnhammer, S. R. Eddy, A. Bateman, R. D. Finn // Nucleic Acids Res. – 2012. – 40. – P.290-301.
12. Thompson J.D. Clustal-W – improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice / J. D. Thompson, D. G. Higgins, T. J. Gibson // Nucleic Acids Res. – 1994. – 22. – P. 4673-4680.
13. Okonechnikov K. Unipro UGENE: a unified bioinformatics toolkit / K. Okonechnikov, O. Golosova, M. Fursov // Bioinformatics. - 2012. - 28. - P. 1166-1167.
14. Blume Ya. B The Plant Cytoskeleton: Key Tool for Agro-Biotechnology / Ya. B. Blume, W. V. Baird, A. I. Yemets, D. Breviario // Springer-Verlag. - 2008. – P. 457.



ПОИСК И АНАЛИЗ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ГЕНОВ АКТИНА В ГЕНОМЕ ЛЬНА

**А. С. Постовойтова, Г. Я. Баер, М. О. Пидюра, Н. Л. Пастухова,
Я. В. Пирко, А. И. Емец, Я. Б. Блюм**

Аннотация. Проведен биоинформационный поиск гомологов генов актина в геноме льна и анализ отобранных последовательностей. Из десяти гомологов генов актина большинство имеет похожую экзон-интронную структуру - преимущественно четыре экзона и три интрона. Кроме того, установлено, что степень идентичности полных последовательностей генов, их кодирующих участков и аминокислотных последовательностей составляет в среднем – соответственно 80 %, 90,5 % и 97 %. Полученные данные свидетельствуют о том, что экзоны относятся к наиболее консервативным участкам генов, а последовательности интронов имеют существенные различия, то есть им присуща высокая степень полиморфизма.

Ключевые слова: лен, актин, локус

SEARCH AND ANALYSIS OF SEQUENCES OF THE ACTIN GENES IN FLAX GENOME

**A. Postovoitova, G. Bayer, N. Pydiura, N. Pastukhova, Ya. Pirko,
A. Yemets, Ya. Blume**

Abstract. The bioinformational search of actin genes homologs and the analysis of selected sequences in the genome of flax were done. Most of ten actin gene homologs have the similar exon-intron structure - four exons and three introns. Also there were found that the degree of identity complete sequences of genes, their coding sites and amino acid sequences on average - 80%, 90,5% and 97% respectively. These data suggest that exons are among the most conserved parts of the genes and introns have significant differences, have a high degree of polymorphism.

Key words: flax, actin, locus