

## Біоінформатичний аналіз гена кукурудзи, що кодує ензим розгалуження крохмалю SBEIIb

Г. І. Сліщук, Т. Ю. Жернаков, Н. Е. Волкова\*

Селекційно-генетичний інститут – Національний центр насіннєзнавства та сортовивчення, вул. Овідіопольська дорога, 3, Одеса, 65036, \*e-mail: [natavolki@ukr.net](mailto:natavolki@ukr.net)

**Мета.** Дослідження поліморфізму гена *ae1* кукурудзи біоінформатичними методами. **Методи.** Глобальне та локальне вирівнювання нуклеотидних та амінокислотних послідовностей, *in silico* трансляція і транскрипція, моделювання транслятивних, дизайн праймерів, філогенетичний аналіз. **Результати.** Проаналізовано 255 нуклеотидних послідовностей гена *ae1* кукурудзи, 500 амінокислотних послідовностей гомологів транслятивних гена *ae1* кукурудзи (гомологів ензиму SBEIIb) і 100 мРНК, що експресуються з гена *ae1* кукурудзи, для встановлення його філогенетичних взаємозв'язків. Біоінформатичними методами досліджено поліморфізм різних ділянок гена *ae1* кукурудзи. Здійснено моделювання ензиму SBEIIb кукурудзи. **Висновки.** За результатами вирівнювання амінокислотних послідовностей гомологів ензиму SBEIIb з'ясовано, що ортологи гена *ae1* присутні лише в однодольних рослин, паралоги – в однодольних, дводольних та інших таксонів, включаючи водорості та тварин. За результатами вирівнювання мРНК рослин, з яких транскрибується ензим SBEIIb, встановлено як ортологи гена *ae1* кукурудзи, так і найближчі паралоги, що кодують ензими розгалуження крохмалю з хлоропластною локалізацією; це свідчить про можливе походження гена *ae1* внаслідок дуплікації гена, що кодує 1,4- $\alpha$ -глюкан-ензим розгалуження крохмалю 2 з хлоропластною або амілопластною локалізацією. В структурі гена *ae1* кукурудзи знайдено регіони, що містять не визначені раніше поліморфні ділянки. До поліморфних регіонів розроблено дизайн праймерів, які дають можливість диференціювати лінії кукурудзи. Визначено, що встановлений поліморфізм теоретично здатний впливати на функцію ензиму та внаслідок цього змінювати концентрацію амілопектину в зерні кукурудзи.

**Ключові слова:** ген *ae1*, ензим розгалуження крохмалю SBEIIb, кукурудза, біоінформатичні підходи.

### Вступ

У біосинтезі крохмалю ензим його розгалуження (starch branching enzyme, SBE) (1,4- $\alpha$ -glucan 6- $\alpha$ -glucosyltransferase; EC 2.4.1.24) є одним з ключових, він формує розгалужену структуру шляхом створення 1,6-зв'язків розгалуження між лінійними ланцюгами. За результатами біохімічних та генетичних досліджень рослинних SBE виявлено три ізоzimні форми – SBEI, SBEII (дві високогомологічні ізоформи SBEIIa та SBEIIb), SBEIII. Встановлено, що вони мають субстратну специфічність. Так, SBEIIa і

SBEIIb віддають перевагу розщепленню і трансферу коротших ланцюгів і каталізують формування 1,6-глікозидних зв'язків, у той час як SBEI віддає перевагу трансферу довгих ланцюгів [1–3].

SBE кукурудзи має три ізоzimні форми – SBEI, SBEIIa, SBEIIb, зокрема SBEI та SBEIIb наявні переважно в ендоспермі, тоді як SBEIIa – в зародку, ендоспермі, листку, інших тканинах. Саме SBEIIb відіграє важливу роль у біосинтезі молекул амілопектину в ендоспермі кукурудзи [4].

Ензим SBEIIb кодується геном *amylose extender (ae1)* (довге плече хромосоми 5), який експресується в ендоспермі з найвищим рівнем серед SBE-генів і відіграє основну роль у детермінації структури амілопектину. Саме SBEIIb має первісне значення для біосинтезу молекул амілопектину в ендоспермі кукурудзи, а не SBEI і SBEIIa: втрата SBEIIa або нестача активності SBEI

Heorhii Slisshchuk  
<http://orcid.org/0000-0003-4245-8557>

Tymofii Zhernakov  
<http://orcid.org/0000-0001-9565-7715>

Nataliia Volkova  
<http://orcid.org/0000-0002-9333-4872>

поодинці не впливає (або має незначний вплив) на склад крохмалю ендосперму та тонкої структури амілопектину. У кукурудзи з рецесивним алелем *ae1* ензим SBEIIb в ендоспермі має низьку активність, що призводить до формування крохмалю з вищим вмістом амілози, більшою кількістю проміжних компонентів і довшим розгалуженим ланцюгом амілопектину, ніж у нормальної кукурудзи з домінуючим геном *ae1*. Вміст амілози в ендоспермі підвищується з нормального середнього рівня 27 до 50% (за деякими джерелами – до 70%) [5].

Кількість і склад крохмалю має важливе значення майже для всіх напрямів використання зерна кукурудзи – від традиційного продовольчого до застосування як відновлюваної сировини для виробництва біопалива. Для поліпшення якості зерна кукурудзи активно використовують біохімічний ефект мутантних генів структури ендосперму, які регулюють біогенез крохмалю та зумовлюють істотний перерозподіл лінійних і розгалужених сополімерів крохмалю. Оскільки найбільший ефект на фракційний склад крохмалю мають мутантні гени *wx1*, *ae1*, *su2*, біоінформатичні та молекулярно-генетичні дослідження генів біосинтезу крохмалю є необхідним етапом розроблення молекулярних маркерів для маркер-супутньої селекції під час створення гібридів кукурудзи для промислового отримання високоякісних крохмалів.

*Мета роботи* – дослідження поліморфізму гена *ae1* кукурудзи біоінформатичними методами для розроблення функціональних маркерів, що будуть використані в селекційних програмах для добору генотипів кукурудзи з певною структурою крохмалю ендосперму зерна.

### Матеріали та методика досліджень

Науково-дослідну роботу виконано у відділі загальної та молекулярної генетики СГП – НЦНС протягом 2015–2016 рр. Матеріалом досліджень були нуклеотидні та амінокислотні послідовності бази даних Національного центру біотехнологічної інформації (National Center for Biotechnology Information, NCBI) [6]. Використано 500 амінокислотних послідовностей транскриптів гомологів гена *ae1* кукурудзи, 100 мРНК – транскриптів гомологів гена *ae1* кукурудзи, 255 нуклеотидних послідовностей гена *ae1* кукурудзи (всі наявні на той час у базі даних NCBI).

Пошук нуклеотидних послідовностей гена *ae1* провадили локальним вирівнюванням за алгоритмом Сміта-Вотермана [7] за допо-

могою он-лайн програми «blastn». Глобальне понуклеотидне та покодонне вирівнювання нуклеотидних послідовностей промотора, екзонів та інтронів здійснювали за алгоритмом Нідлмана-Вунша [8] з використанням програми MEGA5.2 за алгоритмом ClustalW та ClustalW (Codons).

Для *in silico* трансляції, транскрипції та пошуку мікросателітів використовували програму UGENE [9]. Моделювали трансляти *in silico* з використанням автономного сервера гомологічного моделювання структури протеїну SWISS-MODEL [10]. Дизайн праймерів та *in silico* полімеразну ланцюгову реакцію (ПЛР) здійснювали за допомогою програми FastPCR. Умови ПЛР оптимізували згідно з термодинамічними характеристиками матриці та праймерів, обчисленими у програмі. Вірогідність існування побудованих структур обчислювали із застосуванням статистичної суми для будовання матриці вірогідності термодинамічних характеристик системи. Точність передбачених структур за допомогою цієї програми становить 73% для послідовностей розміром до 700 п. н. [11].

Філогенетичний аналіз виконували з використанням пакету MEGA5.2 [12]. Філогенетичні дендрограми реконструювали за методом UPGMA, еволюційні дистанції розраховано з використанням методу Maximum Composite Likelihood, критерієм достовірності був бутстреп тест [13]. Співвідношення синонімічних та несинонімічних заміни аналізували за допомогою функції тестування добору за окремими кодонами програми MEGA5.2.

### Результати досліджень

#### *Філогенетичний аналіз гена ae1 кукурудзи*

Проаналізовано 500 амінокислотних послідовностей гомологів транскриптів гена *ae1* кукурудзи (гомологів ензиму SBEIIb) для встановлення його філогенетичних взаємозв'язків на рівні великих таксонів та 100 мРНК, що експресується з гена *ae1* кукурудзи, – на рівні низьких таксонів.

За результатами вирівнювання 500 амінокислотних послідовностей гомологів ензиму SBEIIb з'ясовано, що ортологи гена *ae1* є лише в однодольних рослин. Паралоги гена *ae1* наявні в однодольних і дводольних та інших таксонів, включаючи водорості та тварин. Паралогами цього гена у тварин є гени, що кодують ензими, які каталізують розгалуження глікогену. Це підтверджує відомий факт, що однодольні, до яких відносять злакові, мають відмінності у ембріогенезі та функціонуванні ендосперму. Завдя-

ки наявності ензимів розгалуження крохмалю утворюється особливий за будовою амілопектин, здатний до компактизації у вигляді неводорозчинних гранул з напівкристалічною структурою, яка є найефективнішою формою зберігання вуглеводів.

За результатами вирівнювання 100 мРНК рослин, з яких транскрибується ензим SBEІІb, встановлено як ортологи гена *ae1* кукурудзи, так і найближчі паралоги, що кодують ензими розгалуження крохмалю з хлоропластною локалізацією. Це свідчить про можливе походження гена *ae1* в результаті дуплікації гена, що кодує 1,4- $\alpha$ -глюкан-ензим розгалуження крохмалю 2 (1,4-alpha-glucan-branching enzyme 2; EC 2.4.1.18) з хлоропластною або амілопластною локалізацією. Продукти експресії предкового гена мали хлоропластну та/або амілопластну локалізацію. Продукти експресії гена *ae1* мають локалізацію в клітинах ендосперму. Ймовірно, що саме ця дуплікація та подальша передислокація продукту цього гена й зумовили еволюційну успішність злакових. Для цього мали відбутися зміни в первинній структурі зазначеного гена – як в промоторній ділянці для регулювання експресії цього гена у зв'язку зі зміною локалізації активності його продукту (в ендоспермі зерна, що розвивається, замість хлоропластів для синтезу амілопектинових гранул для ефективного запасання глюкози, утвореної внаслідок фотосинтезу), так і для забезпечення коректної взаємодії продуктів цього гена з іншими ензимами, що каталізують реакції біосинтезу крохмалю. Також заміни відбулись і в кодуючих ділянках гена для отримання конформації ензиму, ефективною для забезпечення розгалуження амілопектину.

#### *Аналіз поліморфізму гена ae1 кукурудзи*

У базі даних NCBI виявлено 255 нуклеотидних послідовностей, анотованих як ген *ae1* кукурудзи. Зразки гена *ae1* мають середній розмір 23449 п. н. та складаються з промотору, 22 екзонів та 21 інтрона.

255 нуклеотидних послідовностей гена *ae1* кукурудзи проаналізовано з метою виявлення паттернів еволюції. Для промоторної ділянки, представленої 15 нуклеотидними послідовностями, були характерні однонуклеотидні заміни та делеції на фоні значного рівня консервативності.

Регіон екзонів 1–3 був представлений 66 нуклеотидними послідовностями. Тут також були наявні однонуклеотидні заміни та множинні делеції. Взагалі рівень поліморфізму цієї ділянки був невисокий, крім інтрону 2. Щодо характеру еволюції цього гена, то для

нього був характерним стабілізуючий добір, спрямований на збереження первісної амінокислотної послідовності ензиму. Про це свідчать негативні значення за НуPhy аналізом – до -14, це означає, що превалюючими є синонімічні заміни. Можливо, цей регіон має особливе значення для функціонування ензиму та є критичним, тому добір був спрямований на елімінацію несинонімічних замін, які призводять до заміни амінокислот, лише деякі триплети показали позитивне значення, яке за абсолютним значенням не перевищувало негативні показники.

Регіон екзонів 4–8 представляли 17 нуклеотидних послідовностей. Рівень його поліморфізму був високий: наявні численні нуклеотидні заміни й делеції, а також мутація із зсувом рамки зчитування, яка є грубою нуклеотидною мутацією. Для цього регіону був характерним диверсифікуючий добір з превалюванням несинонімічних замін. Це, у поєднанні з наявністю численних поліморфізмів, свідчить про те, що, на відміну від регіону екзонів 1–3, він не відіграє такої значної ролі і, навпаки, еволюція була спрямована на пошук альтернативних варіантів та підвищення поліморфізму.

Регіон екзонів 9–10, представлений 10 нуклеотидними послідовностями, виявився консервативним у зоні екзонів та високополіморфним у зоні інтронів. Для цього регіону був характерним стабілізуючий добір, для 11 сайтів з 12 характерним є превалювання синонімічних замін над несинонімічними. Цей регіон виявляється досить цікавим для подальшого дослідження та пошуку можливих маркерів.

Регіон екзона 11 був представлений 16 висококонсервативними послідовностями. Це означає, що добір був спрямованим на елімінацію майже всіх замін, і свідчить про важливість цієї ділянки гена для функціонування протеїну. Регіон екзонів 12–14, представлений 81 нуклеотидною послідовністю, показав аналогічні дані.

Для регіону екзонів 16–18, представлено му 34 нуклеотидними послідовностями, характерним є високий рівень поліморфізму, встановлено численні однонуклеотидні заміни та делеції, які призводили до зсуву рамки зчитування.

Регіон екзонів 19–22 був висококонсервативним з наявними однонуклеотидними замінами по декількох сайтах.

З огляду на наведене, ймовірно є те, що різні ділянки ензиму SBEІІb мають різне значення для функціонування. Можна також зробити висновок про те, що цей ген



заснав дії доместифікації, що проявилось у вигляді даних, отриманих під час аналізу співвідношення синонімічних та несинонімічних замінів. Враховуючи його очевидну агрономічну важливість, це є цілком зрозумілим.

До найбільш поліморфних ділянок цього гена – інтронів 2, 6 та 8 був розроблений дизайн праймерів (див. табл.). Проведений *in silico* ПЛР-аналіз свідчить, що розроблені пари праймерів дають можливість детектувати поліморфізм у нуклеотидній послідовності гена *ae1*.

Таблиця  
Інформація щодо розроблених пар праймерів до гена *ae1* кукурудзи

Регіон гена	Пара праймерів		Розмір фрагмента ампліфікації, п. н.
	назва	послідовність (5' – 3') (прямого / зворотного)	
Інтрон 2	Zty2	tcctgagggcgagaatgatg ctaaatgggtgacagtaca	261, 264
Інтрон 6	Zty6	ttatttcttctaataataat aaaaatttcccaaacacca	137, 138, 139
Інтрон 8	Zty8	aatggatcatcaatgttat aaaatgatagccgaaaagg	221, 222, 223, 225

#### Моделювання ензиму SBEIIb кукурудзи

Для дослідження впливу поліморфізму гена *ae1* на функціональність ензиму розгалуження крохмалю SBEIIb проведено *in silico* трансляцію проаналізованих нуклеотидних послідовностей. Отримані *in silico* трансляти використано для гомологічного моделювання тривимірної структури ензиму SBEIIb кукурудзи з метою виявлення найперспективніших з погляду впливу на функціонування протеїну сайтів. Прототипом для моделювання структури ензиму слугував аналогічний ензим, що кодується геном у рецесивному стані, який був структурно найближчим гомологом досліджуваного ензиму серед структур, представлених у базі даних Protein data bank (PDB). В результаті аналізу даних виявлено зміну в конформації протеїну. За результатом моделювання знайдено сайт, який, можливо, й зумовлює різницю між доміантними та рецесивними алелями гена *ae1* кукурудзи.

Враховуючи те, що саме ген *ae1* кодує ензим розгалуження крохмалю SBEIIb, що бере участь у перетворенні амілози в амілопектин та результати проведеного дослідження, можна зробити припущення щодо впливу поліморфізму цього гена на співвідношення амілоза/амілопектин, яке визначає напрям використання зерна кукурудзи. Однак для підтвердження цього припущення необхідно проведення експериментальної

верифікації з використанням розроблених пар праймерів у ПЛР *in vitro* на широкій вибірці ліній та гібридів кукурудзи вітчизняної та зарубіжної селекції, що різняться за вмістом амілози та амілопектину. Для встановлення кореляції між алельним станом гена *ae1* та концентрацією амілози/амілопектину необхідно провести біохімічний аналіз зразків насіння кукурудзи та екстраполяцію отриманих даних на результати ПЛР-аналізу.

#### Висновки

За результатами біоінформатичного аналізу гена *ae1* кукурудзи знайдено регіони, що містять раніше не визначені поліморфні ділянки. Для поліморфних регіонів розроблено дизайн праймерів, які дають можливість диференціювати лінії кукурудзи. Визначено, що встановлений поліморфізм теоретично здатний впливати на функцію ензиму розгалуження крохмалю та, як наслідок, змінювати концентрацію амілопектину в зерні кукурудзи. Ці результати є основою розроблення функціональних маркерів гена *ae1*, які можна використовувати в селекційних програмах для добору генотипів кукурудзи з певною структурою крохмалю ендосперму зерна.

#### Використана література

1. Farooq M. Physiology of grain development in cereals / M. Farooq, A. Wahid, K. H. M. Siddique // Handbook of Plant and Crop Physiology / M. Pessaraki (ed.). – 3<sup>rd</sup> ed. – Boca Raton, FL, USA : CRC Press, Taylor & Francis Publishing Group, 2014. – P. 301–308.
2. James M. Seed starch synthesis / M. James, A. Myers // Handbook of Maize: Its Biology / J. L. Bennetzen, S. C. Hake (Eds.). – New York, USA : Springer, 2009. – P. 439–456. doi: 10.1007/978-0-387-79418-1\_22.
3. Seed development: OMICS technologies toward improvement of seed quality and crop yield / G. K. Agrawal, R. Rakwal (Eds.). – Dordrecht, NL : Springer, 2012. – 576 p. doi: 10.1007/978-94-007-4749-4
4. Starch biosynthetic enzymes from developing maize endosperm associate in multisubunit complexes / T. A. Hennen-Bierwagen, F. Liu, R. S. Marsh [et al.] // Plant Physiol. – 2008. – Vol. 146, Iss. 4. – P. 1892–1908. doi: 10.1104/pp.108.116285
5. Hennen-Bierwagen T. A. Genomic specification of starch biosynthesis in maize endosperm / T. A. Hennen-Bierwagen, A. M. Myers // Seed Genomics / P. W. Bewcraft (Ed.). – Hoboken, NJ : John Wiley & Sons, 2013. – P. 123–137. doi: 10.1002/9781118525524.ch7
6. National Center for Biotechnology Information (NCBI) data base [Електронний ресурс]. – Режим доступу : <http://www.ncbi.nlm.nih.gov/>
7. Smith T. F. Identification of common molecular subsequences / T. F. Smith, M. S. Waterman // J. Mol. Biol. – 1981. – Vol. 147, Iss. 1. – P. 195–197. doi: 10.1016/0022-2836(81)90087-5
8. Needleman S. B. A general method applicable to the search for similarities in the amino acid sequence of two proteins / S. B. Needleman, C. D. Wunsch // J. Mol. Biol. – 1970. – Vol. 48, Iss. 3. – P. 443–453.

9. Okonechnikov K. Unipro UGENE: a unified bioinformatics toolkit / K. Okonechnikov, O. Golosova, M. Fursov // *Bioinformatics*. – 2012. – Vol. 28, Iss. 8. – P. 1166–1167. doi: 10.1093/bioinformatics/bts091
  10. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling / K. Arnold, L. Bordoli, F. Kopp, T. Schwede // *Bioinformatics*. – 2006. – Vol. 22, Iss. 2. – P. 195–201. doi: 10.1093/bioinformatics/bti770
  11. Sneath P. H. A. Numerical taxonomy: The principles and practice of numerical classification / P. H. A. Sneath, R. R. Sokal. – San Francisco : W. H. Freeman & Co, 1973. – 573 p.
  12. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods / K. Tamura, D. Peterson, N. Peterson [et al.] // *Mol. Biol. Evol.* – 2011. – Vol. 28, Iss. 10. – P. 2731–2739. doi: 10.1093/molbev/msr121
  13. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism / F. Tajima // *Genetics*. – 1989. – Vol. 123, Iss. 3. – P. 585–595.
1. Farooq, M., Wahid, A., & Siddique, K. H. M. (2014). Physiology of grain development in cereals. In M. Pessarakli (Ed.), *Handbook of Plant and Crop Physiology*. (3<sup>rd</sup> ed.). (pp. 301–308). Boca Raton, FL, USA: CRC Press, Taylor & Francis Publishing Group.
  2. James M., & Myers, A. (2009). Seed starch synthesis. In J. L. Bennetzen & S. C. Hake (Eds), *Handbook of Maize: Its Biology*. (pp. 439–456). New York, USA: Springer. doi: 10.1007/978-0-387-79418-1\_22.
  3. Agrawal, G. K., & Rakwal, R. (Eds). (2012). *Seed development: OMICS technologies toward improvement of seed quality and crop yield*. Dordrecht, NL: Springer. doi: 10.1007/978-94-007-4749-4
  4. Hennen-Bierwagen, T. A, Liu, F., Marsh, R. S., Kim, S., Gan, Q., Tetlow, I. J., ... Myers, A. M. (2008). Starch biosynthetic enzymes from developing maize endosperm associate in multi-subunit complexes. *Plant Physiol.*, 146(4), 1892–1908. doi: 10.1104/pp.108.116285
  5. Hennen-Bierwagen, T. A., & Myers, A. M. (2013). Genomic specification of starch biosynthesis in maize endosperm. In P. W. Becraft (Ed.), *Seed Genomics*. (pp. 123–137). Hoboken, NJ: John Wiley & Sons. doi: 10.1002/9781118525524.ch7
  6. *National Center for Biotechnology Information (NCBI) data base*. (n.d.). Retrieved from <http://www.ncbi.nlm.nih.gov>
  7. Smith, T. F., & Waterman, M. S. (1981). Identification of common molecular subsequences. *J. Mol. Biol.*, 147(1), 195–197. doi: 10.1016/0022-2836(81)90087-5
  8. Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, 48(3), 443–453.
  9. Okonechnikov, K., Golosova, O., & Fursov, M. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 28(8), 1166–1167. doi: 10.1093/bioinformatics/bts091
  10. Arnold, K., Bordoli, L., Kopp, F., & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling. *Bioinformatics*, 22(2), 195–201. doi: 10.1093/bioinformatics/bti770
  11. Sneath, P. H. A., & Sokal, R. R. (1973). *Numerical taxonomy: The principles and practice of numerical classification*. San Francisco: W. H. Freeman & Co.
  12. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, 28(10), 2731–2739. doi: 10.1093/molbev/msr121
  13. Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), 585–595.

## References

УДК 633.15.631.527

**Слищук Г. И., Жернаков Т. Ю., Волкова Н. Е.\*** Биоинформатический анализ гена кукурузы, кодирующего энзим ветвления крахмала SBEIIb // Сортовивчення та охорона прав на сорти рослин. – 2016. – № 3. – С. 13–18. [http://dx.doi.org/10.21498/2518-1017.3\(32\).2016.75972](http://dx.doi.org/10.21498/2518-1017.3(32).2016.75972)

Селекційно-генетический інститут – Національний центр семеноведення і сортознучення, ул. Овідіопольська дорога, 3, Одеса, 65036, Україна, \*e-mail: [natavolki@ukr.net](mailto:natavolki@ukr.net)

**Цель.** Исследование полиморфизма гена *ae1* кукурузы биоинформатическими методами. **Методы.** Глобальное и локальное выравнивание нуклеотидных и аминокислотных последовательностей, *in silico* трансляция и транскрипция, моделирование транслятов, дизайн праймеров, филогенетический анализ. **Результаты.** Проанализированы 255 нуклеотидных последовательностей гена *ae1* кукурузы, 500 аминокислотных последовательностей гомологов транслятов гена *ae1* кукурузы (гомологов энзима SBEIIb) и 100 мРНК, экспрессирующихся с гена *ae1* кукурузы, для установления его филогенетических взаимосвязей. Биоинформатическими методами исследован полиморфизм различных участков гена *ae1* кукурузы. Осуществлено моделирование фермента SBEIIb кукурузы. **Выводы.** По результатам выравнивания аминокислотных последовательностей гомологов фермента SBEIIb установлено, что ортологи гена *ae1* присутствуют лишь у однодольных растений, паралоги – у однодольных, двудольных и других таксонов,

включая водоросли и животных. По результатам выравнивания мРНК растений, с которых транслируется энзим SBEIIb, определены как ортологи гена *ae1* кукурузы, так и ближайшие паралоги, кодирующие ферменты ветвления крахмала с хлоропластной локализацией; это свидетельствует о возможном происхождении гена *ae1* в результате дубликации гена, кодирующего 1,4- $\alpha$ -глюкан-энзим ветвления крахмала 2 с хлоропластной или амилопластной локализацией. В структуре гена *ae1* кукурузы найдены регионы, которые включают в себя не определенные ранее полиморфные участки. К полиморфным регионам разработан дизайн праймеров, которые позволяют дифференцировать линии кукурузы. Определено, что обнаруженный полиморфизм теоретически способен влиять на функцию фермента и в результате этого изменять концентрацию амилопектина в зерне кукурузы.

**Ключевые слова:** ген *ae1*, энзим ветвления крахмала SBEIIb, кукуруза, биоинформатические подходы.

UDC 633.15.631.527

**Slisichuk, G. I., Zhernakov, T. Yu., & Volkova, N. E.\*** (2016). Bioinformatic analysis of maize gene encoding starch branching enzyme SBEIIb. *Sortovivčennâ ohor. prav sorti roslin* [Plant Varieties Studying and Protection], 3, 13–18. [http://dx.doi.org/10.21498/2518-1017.3\(32\).2016.75972](http://dx.doi.org/10.21498/2518-1017.3(32).2016.75972)

*Plant Breeding and Genetics Institute – National center of Seed and Cultivar Investigation, 3, Ovidiopolska Doroga st., Odesa, 65036, Ukraine, \*e-mail: natavolki@ukr.net*

**Purpose.** Investigation of maize *ae1* gene polymorphism by bioinformatic methods. **Methods.** Global and local alignment of the nucleotide and amino acid sequences, *in silico* translation and transcription, translates modeling, primers design, phylogenetic analysis. **Results.** 255 nucleotide sequences of maize *ae1* gene, 500 amino acid sequences of homology translates of maize *ae1* gene (SBEIIb enzyme homologs) and 100 mRNA expressed from the maize *ae1* gene were analyzed to establish phylogenetic relationships. Polymorphism of maize *ae1* gene different regions was investigated by bioinformatic methods. Modeling of the maize enzyme SBEIIb was performed. **Conclusions.** According to the results of amino acid sequences of SBEIIb enzyme homologs alignment, it was found that *ae1* gene orthologs are present only in monocots, paralogs – in monocots, dicots, and other taxa, including algae and animals. Based on the results

of alignment of plants mRNA from which enzyme SBEIIb is translated, maize *ae1* gene orthologs and the nearest paralogs encoding starch branching enzymes with chloroplast localization were defined; this suggests a possible origin of *ae1* gene due to duplication of the gene encoding the 1,4-alpha-glucan-branching enzyme 2 with chloroplast or amyloplast localization. In the maize *ae1* gene structure, regions were found that include polymorphic sites not defined previously. For the polymorphic sites design primers were developed that allowed to differentiate the maize lines. It was determined that the detection of polymorphism in theory can influence the enzyme function and, as a result, change the concentration of amylopectin in maize grain.

**Keywords:** *ae1* gene, starch branching enzyme SBEIIb, maize, bioinformatic.

Надійшла 23.05.2016