

## Біоінформатичний аналіз нуклеотидних послідовностей гена ацетолатат синтази (*als*) різних представників родини Злакових

А. В. Кирієнко<sup>1,2\*</sup>, Н. Л. Щербак<sup>1</sup>, Ю. В. Симоненко<sup>1,2</sup>

<sup>1</sup>Інститут клітинної біології та генетичної інженерії НАН України, вул. Академіка Заболотного, 148б, м. Київ, 03143, Україна, \*e-mail: [anastasija.kirienko@gmail.com](mailto:anastasija.kirienko@gmail.com)

<sup>2</sup>Всеукраїнський науковий інститут селекції, вул. Васильківська, 30, м. Київ, 03022, Україна

**Мета.** Провести біоінформатичний аналіз та порівняти цільові ділянки гена ацетолатат синтази (*als*) у декількох представників родини Злакових і на основі отриманих даних дослідити можливість створення уніфікованої генетичної конструкції для направленої редагування гена *als* за допомогою системи CRISPR-Cas9. **Методи.** Сиквенс гена *als* різних представників родини Злакових було отримано з бази даних NCBI: Nucleotide. Для порівняння було використано фрагмент гена *imi-2* пшениці м'якої лінії 'TealIMI11A' у двох ділянках сиквенсу: 367–390 та 1729–1749 нуклеотидів. Для оцінювання наявності нуклеотидних замін в робочих сиквенсах гена *imi-2* використовували інструмент "Sequence Viewer 3.37.0". Дендрограму будували з використанням інструменту 'Blast Tree' з ресурсу NCBI: Blast: Nucleotide. **Результати.** Було проаналізовано нуклеотидні послідовності сімох різних видів: пшениці м'якої (*Triticum aestivum* L.), вівсюга звичайного (*Avena fatua* L.), ячменю звичайного (*Hordeum vulgare* L.), рису посівного (*Oryza sativa* L.), кукурудзи (*Zea mays* L.), сорго алепського (*Sorghum halepense* Pers.) та егілопса Тауша (*Aegilops tauschii* Coss.). На основі порівняльного аналізу сиквенсу ділянок гена *imi-2* для 7-ми генотипів було побудовано філогенетичне дерево, яке показало, що досліджені види можна поділити на два блоки. До одного з них увійшли кукурудза та сорго алепське, а до другого – рис посівний, вівсюг звичайний, ячмінь звичайний, пшениця м'яка та егілопс Тауша. Визначення ступеня гомології між послідовністю 367–390 нуклеотиду пшениці м'якої та іншими видами показало, що абсолютною була гомологія з відповідними послідовностями рису посівного, егілопса Тауша та вівсюга звичайного (100%). Найменшою нуклеотидна спорідненість виявилась для кукурудзи та сорго алепського – 83,3%. У ділянці 1729–1749 нуклеотидів гена *imi-2* жоден з 6 сиквенсів не показав 100% гомології з послідовністю пшениці м'якої. Найвищою вона була до ячменю звичайного та егілопсу Тауша – 95,2%, а найменшою для рису посівного, кукурудзи та сорго алепського – по 80,9%. **Висновки.** Проведений аналіз підтверджує значний ступінь гомології послідовності гена *als* для різних видів родини Злакових. Це дозволяє зробити припущення про можливість створення універсальної генетичної конструкції, за допомогою якої можна редагувати геном та отримувати рослини, стійкі до гербіцидів різних представників цієї родини. Зважаючи на вищий ступінь гомології послідовностей для таких видів як пшениця м'яка, егілопс Тауша, ячмінь звичайний, рис посівний та вівсюг звичайний, можна припустити, що ефективною відповідною генетичною конструкцією буде для редагування гена *als* саме цих генотипів.

**Ключові слова:** пшениця м'яка; стійкість до гербіцидів; ген *als*; редагування генома.

### Вступ

Спонтанне формування стійкості до гербіцидів серед дикорослих бур'янів є розповсюдженим явищем і часто зустрічається в при-

родних умовах. Поява толерантності до дії гербіцидів обумовлена точковими мутаціями, що виникають у ключових для резистентності генах рослини. Резистентність до гербіцидів є еволюційним процесом і залежить від генетичних факторів, серед яких перш за все можна виділити кількість резистентних генів, їхні частоти та механізм успадкування [1].

Досить розповсюдженим типом набутої толерантності рослин до дії гербіцидів є перехресна резистентність. Вона виникає коли

Anastasiia Kyriienko

<https://orcid.org/0000-0002-8117-5288>

Natalia Shcherbak

<https://orcid.org/0000-0002-2478-8408>

Yuri Symonenko

<https://orcid.org/0000-0002-5597-3315>

рослина формує стійкість до декількох гербіцидів, що належать до одного класу. Так, якщо на певній території популяція бур'янів спрямовано піддавалась дії лише гербіциду імазамоксу, то з часом деякі дикорослі рослини із цієї популяції набудуть резистентності як до імазамоксу, так й до імазатопіру та інших гербіцидів класу імідазолінонів [1].

Гербіциди групи імідазолінів, сульфонілсечовини, триазолпіримідинів, піримідилоксисбензоатів та сульфоніламінокарбоксилтриазолінів інгібують роботу ензиму ALS (Acetolactate synthase) – ацетолактат синтази, також відомої як AHAS (Acetohydroxyacid synthase) – ацетогідроксил синтаза [2]. За своєю будовою цей ензим є октамером, який складається із 4 каталітичних та 4 регуляторних субодиниць (рис. 1). За своєю будовою каталітична субодиниця є димером, що складається із двох мономерів, кожен з яких розміром близько 670 амінокислотних залишків та має у своєму складі три домени –  $\alpha$ ,  $\beta$  та  $\gamma$ . У ензиму ALS є два активних сайти зв'язування із субстратом, які утворені амінокислотами із обох мономерів каталітичної субодиниці [3]. Імідазолінон-резистентні рослини містять мутації, що обумовлюють стійкість до гербіциду, у будь-якому з трьох доменів:  $\alpha$  та  $\gamma$  [4], або  $\beta$  домені [5]. До деяких таких мутацій в гені *als* *Saccharomyces cerevisiae* відносять заміни в домені  $\alpha$  таких амінокислот як Gly116, Ala117, Leu119, Pro192, Ser194, Ala200 та Lys251; у ділянці  $\beta$  домену – Met354 та Asp379; у складі  $\gamma$  домену – Val583, Trp586, Phe590 та Gly657 [5, 6].

Ензим ALS здійснює каталіз двох паралельних біохімічних реакцій, що пов'язані

із біосинтезом амінокислотного ланцюга. Під час першої реакції відбувається конденсація двох молекул пірувату, що згодом веде до утворення валіну та лейцину. В іншій реакції ALS каталізує конденсацію молекули пірувату з альфа-кетобутиратом, що веде до утворення ізолейцину [7].

Молекула гербіциду здатна зв'язуватись із субстрат-зв'язуючою ділянкою ензиму ALS, що призводить до блокування центрів зв'язування із субстратом і гальмує синтез відповідних амінокислот в рослині. Як наслідок, знижується білковий синтез та число поділів рослинних клітин. Найуразливішими є меристематичні тканини, в яких процеси білкового синтезу відбуваються найактивніше. Зрештою рослина починає сповільнювати темпи росту і гине [7].

Відомо, що у пшениці м'якої (*Triticum aestivum* L.) стійкість до гербіцидів імідазолінового ряду виникає унаслідок точкової мутації в активному сайті гена *als*, що зрештою веде до утворення ізоензиму резистентного до гербіцидів [8]. Загалом відомо 6 найбільше досліджених мутацій, які пов'язанні із амінокислотою заміною в позиціях Ala122, Pro197, Ala205, Arg377, Trp574, Ser653 (назва і порядок амінокислот відповідають *Arabidopsis thaliana* L.) [8]. Амінокислотні залишки Ala122, Pro197, Ala205 розміщені ближче до N-кінця ензиму, а амінокислоти Asp376, Trp574, Ser653 ближче до C-кінця. До прикладу, заміна проліну на серин або аланін у 197 амінокислотній позиції призводить до набуття рослиною резистентності до гербіцидів групи сульфонілсечовини; аланіну 122 та серину 653 до імідазолінів; триптофану 574 до імідазолінів, сульфо-

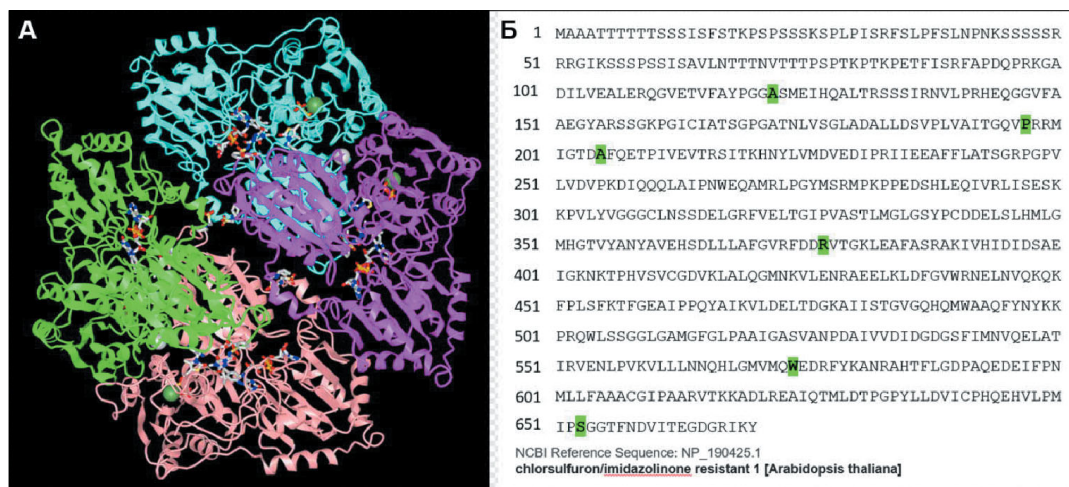


Рис. 1. Білок ALS *Arabidopsis thaliana* L.: А) 3D модель структури білка (взято з бази даних NCBI-Protein Structure [www.ncbi.nlm.nih.gov/Structure](http://www.ncbi.nlm.nih.gov/Structure)); Б) амінокислотна послідовність білка ALS, зеленим кольором позначено амінокислоти, заміна яких асоціюється із стійкістю рослини до імідазолінів

нілсечовини та триазолпіримідинів; натомість заміна аланіну 205 веде до резистентності проти всіх гербіцидів, які здатні зв'язуватись із ацетолактат синтазою [9]. Варто зауважити, що мутація із заміною аланіну 122 на треонін чинить вираженіший ефект резистентності до гербіцидів, ніж мутації, пов'язані із серином 653. Досить поширеною є заміна серину 653 на аспарагін або на треонін, що зумовлює стійкість до імідазолінонів. Це є наслідком заміни коду AGC на ACC, одноклеточної заміни гуаніну на цитозин в 1742 позиції для генотипу пшениці м'якої 'CDC Teal' [7, 9]. Крім того, амінокислотні заміни Asp376 та Ser654 також характерні для пшениці м'якої. На сьогодні створено сорти пшениці твердої (*Triticum durum* L.) – 'Rafalin' та 'Antonin', стійкі до імідазолінів, якраз на основі заміни амінокислоти серин на аспарагін у 653 позиції. Так, для генотипу 'Antonin' мутація стосується лише гена *imi-2*, а для 'Rafalin' генів *imi-1* та *imi-2*. Ці генотипи є найстійкішими до імазамоксу, імазоквіну та імазапіру. Генотип 'Rafalin' походить від схрещування не стійкого сорту 'Gazul' із стійкою комерційно зареєстрованою формою 'Pantera Clearfiel'®, остання містить мутацію у генах *imi-1* та *imi-2*. Генотип твердої пшениці 'Antonin' походить від схрещування сорту 'Simeto' з *Triticum durum*, який містить мутацію у гені *imi-2* і є наслідком схрещування *Triticum turgidum* L. з *Hordeum chilense* Roem. & Schult [9].

У пшениці м'якої гени *als* із мутаціями стійкості до імідазолінів представлені у кожному геномі (геноми гексаплоїдної пшениці А, В та D) та називаються *imi-1*, *imi-2* та *imi-3* [10], також відомі як *ahasL-B1*, *ahasL-D1* та *ahasL-A1* [9]. У геномі гексаплоїдної пшениці ці гени локалізовані на довгому плечі хромосом 6A, 6D та 6B. Мутацію в 6DL називають *Imi-1* [4]. Відомо, що у пшениці м'якої генетичні мутації, які надають стійкості до імазамоксу є ядерними, а не цитоплазматичними, та успадковуються як домінантні [10], або напів домінантні [5, 11].

На сьогодні для пшениці м'якої в базі даних NCBI: Nucleotide можна знайти інформацію про сиквенс частини гена з корисною мутацією *imi-2*, який також називають *TealIMI11A*. Загалом 'Teal' (CDC Teal) – це назва сорту ярої пшениці м'якої, зерно якої було оброблено етил метан сульфонатом. Згодом отримані M2 рослини були оброблені імазамоксом. З них 6 ліній показали високий рівень толерантності до цього гербіциду, а саме лінії 1A, 9A, 11A, 15A та 16A. Тому

напис 'TealIMI11A' означає, що це 11A лінія сорту 'Teal' ярої пшениці м'якої, яка толерантна до імідазолінонів [4, 10].

На сьогодні на ринку зареєстровано велику кількість імідазолінових гербіцидів з такими комерційними назвами, як: PURSUIT® (активна речовина імазетапір), CARDE® (імазанік), RAPTOR® (імазамокс), SCEPTER® (імазаквін), ASSERT® (імазетобенз), ARSENAL® (імазапір), ODYSSEY® (імазапір та імазамокс) та EURO-LIGHTING® (імазапір та імазамокс) [12]. Також уже створено велику кількість рослинних форм стійких до імідазолінонів, серед яких є кукурудза *Zea mays* L. (1992 рік), ріпак *Brassica napus* L. (1995 рік), гексаплоїдна пшениця *T. aestivum* L. (2001 рік), рис *Oryza sativa* L. (2003 рік) та соняшник *Helianthus annuus* L. (2003 рік) [4].

Мутації, пов'язані із заміною амінокислот, можуть бути, як спонтанними і нести випадковий характер, так і штучно індукованими. Відбір та селекція зразків пшениці м'якої та інших злакових, що несуть природні мутації гена *als* може бути достатньо складним процесом та витратним за часом, а отриманий результат може бути мінімальним. Тому для створення рослинних форм стійких до гербіцидів імідазолінового ряду та деяких інших доцільним виглядає використання штучно індукованих мутацій. У даному випадку можливі два основні і найпоширеніші варіанти. Перший передбачає використання хімічного мутагену, наприклад азиду натрію [13, 14], або 5-бромоурацилу [15]. Насіння покоління M0, оброблене цими агентами, висівають у закритий ґрунт, після чого здійснюють селекцію насіння M1 на предмет стійкості до певного гербіциду. Відібрані стійкі генотипи проходять молекулярно-генетичний скринінг та підтвердження наявності корисної мутації, що надала стійкості. Цей підхід підвищує ймовірність отримання потрібного генотипу, однак має такі недоліки, як необхідність використання сильних мутагенів та випадковість отриманих мутацій, що можуть виникнути і в інших генах рослини.

Сучаснішим підходом зі створення стійких до гербіцидів рослин є спрямоване редагування генома за допомогою системи CRISPR-Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats). Ідея цього методу полягає в тому, що, володіючи інформацією про нуклеотидну послідовність певного гена, за допомогою нуклеази Cas9 можна цілеспрямовано внести подвійний розрив ДНК у конкретну ділянку. Після штучного внесення такого ушкодження репарація подвійного розриву ДНК в клітині відбувається за

допомогою негомологічного з'єднання кінців (Non-homologous end joining, NHEJ), що зазвичай призводить до втрати або заміни певної кількості нуклеотидів у місці розриву. Для реалізації такого підходу має бути створений генетичний вектор, що містить ген нуклеази Cas9 та послідовність комплементарної до ділянки інтересу guide-РНК, завдяки якій відбувається розпізнавання послідовності у потрібному місці та внесення випадкових змін (наприклад місенс або нонсенс мутацій тощо) [16].

Аналіз та порівняння між собою генетичних послідовностей гена *als* різних представників родини Злакових допоможе систематизувати та узагальнити інформацію, щодо мутацій у гені ацетолактат синтази. Це дає підґрунтя для створення генетичних конструкцій для CRISPR-Cas9 генетичного редагування генома промислово важливих злакових культур (пшениця, ячмінь, овес) для створення нових ліній, стійких до гербіцидів імідазолінового ряду. Представники родини Злакових (*Poaceae*) можуть мати ділянки ДНК з високим ступенем гомології. Підтвердження цього припущення дозволить оптимізувати дизайн генетичної конструкції для редагування генома злакових та створити універсальний вектор, який підходить одразу до декількох видів. Наприклад, РАМ (Protospacer adjacent motif) – суміжний мотив протоспейсера, ділянка, яка розпізнається нуклеазою Cas9 і розміщена поруч із послідовністю гена [16]. Якщо для декількох видів у РАМ ділянках послідовностей ключового гена будуть відсутні мутації, це дозволить використовувати одну й ту саму конструкцію для редагування цього гена у різних видів.

**Мета досліджень** – здійснити біоінформатичний аналіз та порівняння цільових ділянок гена ацетолактат синтази (*als*) у декількох представників родини Злакових і на основі отриманих даних дослідити можливість створення уніфікованої генетичної конструкції для направленої редагування гена *als* за допомогою системи CRISPR-Cas9.

## Матеріали та методика досліджень

Сиквенс гена *als* представників родини Злакових (*Poaceae*) було отримано із бази даних NCBI: Nucleotide. Порівнювали ці послідовності за допомогою ресурсу NCBI: Blast: Nucleotide (blast.ncbi.nlm.nih.gov). Ключовою наявною інформацією про нуклеотидні послідовності гена *als*, в роботу було залучено 7 генотипів від різних представників родини *Poaceae*. Детальнішу інформацію про кожен генотип представлено у таблиці 1.

Для роботи обрали представників родини Злакових, які, за винятком егілопса Тауша та сорго алепського, є важливими сільськогосподарськими культурами. Егілопс Тауша не має господарського значення, однак йому властива тісна еволюційна генетична спорідненість з пшеницею м'якою.

Аналіз проводили із залученням часткового сиквенсу гена *imi-2*, довжиною 1797 нуклеотидів, лінії пшениці м'якої 'TeaIMI11A' і здійснювали його порівняння із 6-ма іншими сиквенсами різних представників родини Злакових. Порівняння виконували у двох ділянках сиквенсу розміром 21 нуклеотид, що вміщує 7 кодонів: 367–390 (послідовність 1) та 1729–1749 (послідовність 2). Звертали увагу саме на ці регіони, оскільки з літературних джерел відомо, що вони можуть бути придатними для внесення нуклеотидних змін [17], а отже, і редагування генома пшениці м'якої за технологією CRISPR-Cas9. Послідовність 1 містить 8 кодонів у межах 24 нуклеотидів, а послідовність 2 – 7 кодонів у межах 21 нуклеотиду.

На основі проаналізованих сиквенсів 7 представників родини Злакових було побудовано дендрограму за допомогою інструменту 'Blast Tree' з ресурсу NCBI: Blast: Nucleotide. Побудову дерева виконували за допомогою 'Fast Minimum Evolution' та коефіцієнтом максимальної різниці сиквенсу – 0,75. Для порівняння нуклеотидних послідовностей використали 'Sequence Viewer 3.37.0'.

Також було проведено оцінювання ступеня гомології проаналізованих послідовнос-

Таблиця 1

### Інформація про сиквенс гена *imi-2* представників родини Злакових

№ з/п	Українська назва	Латинська назва	Назва сорту/форми/ізоляту	Номер послідовності гена в GenBank NCBI
1	Пшениця м'яка	<i>Triticum aestivum</i> L.	TeaIMI11A, imi2	AY210408.1
2	Ячмінь звичайний	<i>Hordeum vulgare</i> L.	Golden Promise	LT601589.1
3	Рис посівний	<i>Oryza sativa</i> L.	Japonica Group	AY885674.1
4	Егілопс Тауша	<i>Aegilops tauschii</i> Coss.	BJ-S	FJ997628.1
5	Вівсюг звичайний	<i>Avena fatua</i> L.	–	FJ997632.1
6	Кукурудза	<i>Zea mays</i> L.	–	X63553.1
7	Сорго алепське	<i>Sorghum halepense</i> Pers.	Біотип D	KJ538786.1



тей між собою. Гомологію визначали за формулою:

$$G = (m/n) \times 100,$$

де:  $m$  – кількість однакових нуклеотидів,  $n$  – загальна кількість нуклеотидів. Отриману величину позначали у відсотках. Відповідно повну гомологію, де дві нуклеотидні послідовності повністю відповідають одна одній, позначали як 100%. Натомість чим меншим був цей показник, тим нижчою була гомологія. Загальний ступінь гомології для всієї послідовності гена *als* брали із NCBI: Blast: Nucleotide для порівняння окремих сиквенсів.

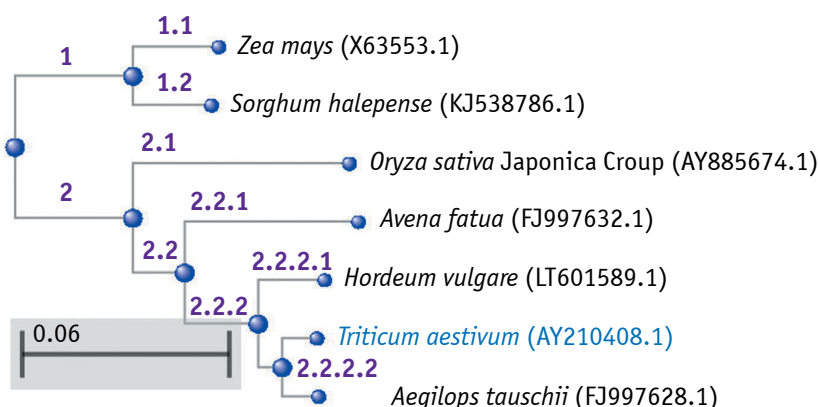


Рис. 2. Дендрограма філогенетичних зв'язків 7 представників родини Злакових на основі сиквенсів гена *als* (*imi-2*): блакитним кольором позначено генотип пшениці м'якої, відносно якого здійснювали порівняння

Для 6 сиквенсів гена *imi2* було здійснено оцінювання рівня гомології відносно сиквенсу пшениці м'якої (табл. 2). Так, послідовність 1 (367–390 нуклеотидів) пшениці м'якої демонструвала найвищу гомологію на рівні 100% з відповідними послідовностями рису посівного, егілопса Тауша та вісюга звичайного. До ячменю звичайного гомологія була трохи меншою – 95,8%. Натомість найменший ступінь гомології у пшениці для проаналізованої ділянки виявився відносно кукурудзи та сорго алепського – по 83,3%.

Оцінювання нуклеотидної послідовності 1729–1749 показало відсутність повної гомології на рівні 100%. Натомість найвищою гомологія була до ячменю звичайного та егілопса Тауша – 95,2%, дещо меншою до вісюга звичайного – 85,7% та найменшою для рису посівного, кукурудзи та сорго алепського – по 80,9% (табл. 2).

Також оцінювали загальний рівень гомології для усієї послідовності розміром 1797 нуклеотидів. Виявилось, що найвищою гомологія ділянок гена пшениці була до егілопса Тауша – 97,9%. Далі – до ячменю зви-

## Результати досліджень

Унаслідок біоінформатичного аналізу нуклеотидних послідовностей гена *als* було побудовано дендрограму досліджуваних представників родини Злакових (рис. 2). Так, було виокремлено два блоки: 1 та 2. До блоку 1 увійшли кукурудза (гілка 1.1) та сорго алепське (гілка 1.2). До другого блоку окремою гілкою 2.1 увійшов рис посівний та група 2.2. До останньої увійшли вісюг звичайний (гілка 2.2.1) та окремо підгрупа, що поєднала ячмінь звичайний (гілка 2.2.2.1), пшеницю м'яку та егілопс Тауша.

чайного – 96,4%, вісюга звичайного – 91,5%, рису посівного – 89,1%. Найменший рівень гомології виявлено при порівнянні послідовності пшениці із сорго алепським – 86,7% та кукурудзою – 86,4% (табл. 2). Загалом аналіз показав, що егілопс Тауша під час порівняння окремих ділянок сиквенсу та повної послідовності гена, демонстрував найвищий ступінь гомології до пшениці м'якої. Це можна пояснити тим, що в еволюційному плані пшениця м'яка та егілопс Тауша досить близькі. Відомо, що геном амфідиплоїдної пшениці м'якої представлений як AABBDD. Вважають, що DD має безпосереднє походження від дикорослого злаку егілопса Тауша. Припускають, що геном сучасної пшениці м'якої AABBDD утворився унаслідок гібридизації між тетраплоїдним видом *Triticum turgidum* L. (AABB) та *Aegilops tauschii* Coss. (DD) [18, 19].

Сиквенс гена *imi2* було проаналізовано на наявність конкретних нуклеотидних замін. Спершу розглянули нуклеотидну послідовність 367–390 (рис. 3). На думку Zhang R. і співавторів у цьому регіоні можна робити ну-

Таблиця 2

Оцінювання ступеня гомології проаналізованих послідовностей гена *als* для представників родини Poaceae

№ з/п	Генотип	Гомологія, %		Загальний рівень гомології, %
		367–390	1729–1749	
1	Пшениця м'яка	100	100	100
2	Ячмінь звичайний	95,8	95,2	96,4
3	Рис посівний	100	80,9	89,1
4	Егілопс Тауша	100	95,2	97,9
5	Вівсюг звичайний	100	85,7	91,5
6	Кукурудза	83,3	80,9	86,4
7	Сорго алепське	83,3	80,9	86,7

клеотидні заміни [17]. Ними показано, що мутації у вигляді нуклеотидної заміни в позиції 373, 374, 375 та 376 найвірогідніші і спричиняють стійкість до гербіцидів імідазолінового ряду. Заміна цитозину 373 на тимін в кодоні

CCC (на TCC) веде до утворення Ser174 замість Pro174. Заміна першого цитозину в цьому кодоні на гуанін (GCC) веде до утворення Ala174. Одночасна заміна першого та другого, а також усіх трьох цитозинів на тимін (TTC/TTT) веде до утворення Phe174. Також, якщо в першому цитозині сусіднього кодону CGC відбувається заміна на тимін (TGC), то замість Agr175 утворюється Cys175 [17].

У дослідженні в нуклеотидній послідовності ячменю звичайного було виявлено заміну кодону CCC на CCA. Також спостерігали чотири однакові заміни для послідовностей кукурудзи та сорго алепського, а саме заміна GTC на GTG, та CCC на CCG, та CGC на CGA. Остання нуклеотидна заміна припадала на ділянку PAM – заміна ATC на ATT. Так, цитозин в цьому кодоні входить до складу PAM ділянки.

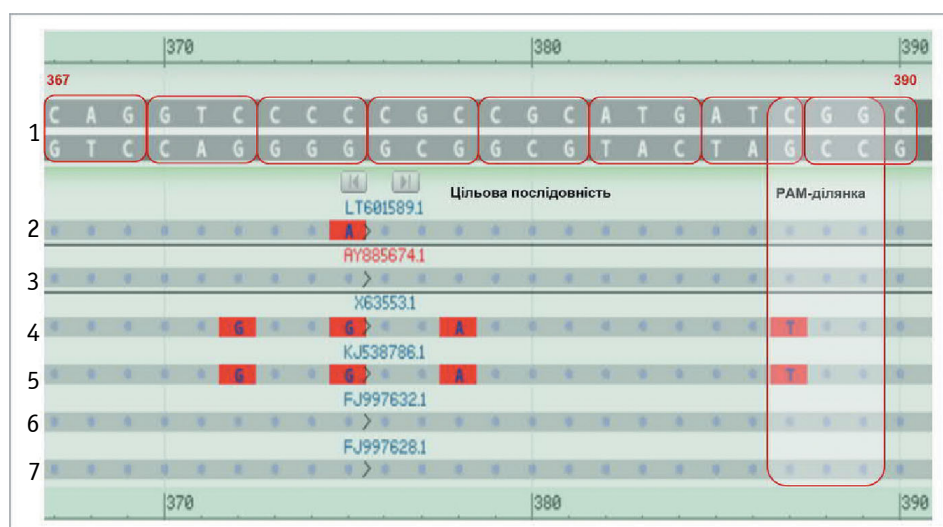


Рис. 3. Ділянка 367–390 нуклеотидної послідовності гена *imi-2* для представників родини Злакових: 1) *Triticum aestivum* L., 2) *Hordeum vulgare* L., 3) *Oryza sativa* L., 4) *Zea mays* L., 5) *Sorghum halepense* Pers., 6) *Avena fatua* L., 7) *Aegilops tauschii* Coss.

Для редагування гена *als* даного регіону можна створити універсальну конструкцію для пшениці м'якої, рису посівного, вівсюга звичайного, егілопса Тауша та ячменю звичайного. Це пов'язано з тим, що проаналізовані ділянки не містили суттєвих відмінностей.

Також було проаналізовано представлену послідовність у межах 1729–1749 нуклеотидів (рис. 4). На думку Zhang R. та співавторів одно- та двонуклеотидні заміни, які виникають у двох послідовних кодонах GGT (регіон 1744–1749) мають найбільше відношення до стійкості до імідазолінонів [17]. Також відомо, що для вівсюга звичайного заміна Ser653Thr та Ser653Asn дає стійкість до гербіцидів імідазолінового ряду [20]. Крім

того, у рису посівного в гені *ahas* однонуклеотидна заміна G на A в 1880 позиції призводить до заміни серину на аспарагін у 627 позиції амінокислотного ланцюга [21]. Також відомо, що амінокислотна заміна гліцину на глутамат у 628 позиції у рису підвищує стійкість до імідазолінів, однак у 4,9 разів нижча, ніж при заміні серину на аспарагін у 627 позиції [22].

Виявилось, що в проаналізованій ділянці представлено два кодони GGT. У рису посівного спостерігалась заміна T на G, а в другому кодоні T на C. У кукурудзи та у сорго алепського зміни стосувались лише одного GGT кодону, де T замінюють на G. Для всіх розглянутих нуклеотидних послідовностей PAM ділянка була представлена коконом

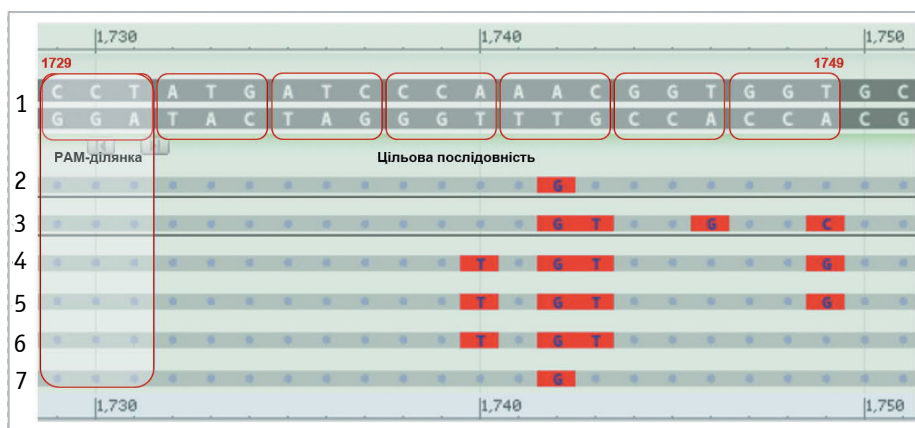


Рис. 4. Ділянка 1729–1749 нуклеотидної послідовності гена *imi-2* для представників родини Злакових: 1) *Triticum aestivum* L., 2) *Hordeum vulgare* L., 3) *Oryza sativa* L., 4) *Zea mays* L., 5) *Sorghum halepense* Pers., 6) *Avena fatua* L., 7) *Aegilops tauschii* Coss.

ССТ. Це означає, що для усіх проаналізованих генотипів можна створити універсальну генетичну конструкцію, за допомогою якої, можна здійснювати редагування гена *als*.

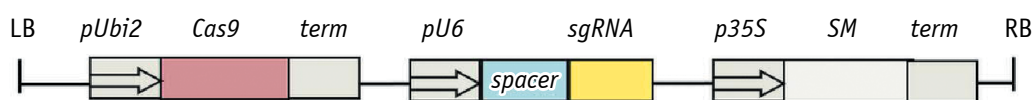


Рис. 5. Схематична будова генетичного вектора для редагування гена *als* представників родини Злакових: term – термінатор транскрипції відповідного гена, SM – ген селективного маркера, p35S, pUbi2, pU6 – промотори, sgRNA – одинарна направляюча РНК, Cas9 – ген нуклеази, spacer – спейсерна ділянка

Генетична конструкція для направленої редагування гена *als* Злакових має містити кілька ключових елементів (рис. 5). Насамперед ген нуклеази Cas9 та послідовність направляючої РНК (guide RNA) з ділянкою спейсер, яка є комплементарною послідовності гена фермента AHAS (ALS) і своєрідним «шаблоном» для розпізнавання відповідного сайту в геномі рослини. Наші дослідження були спрямовані на виявлення саме цієї послідовності, спільної для ряду представників родини Злакових. Направляюча РНК з «універсальним» спейсером зможе «знаходити» та забезпечувати редагування гена *als* у декількох видів.

### Висновки

Для генетичного редагування ділянки 367–390 нуклеотидів гена *als* можна створити універсальну конструкцію для 5 видів злакових: пшениці м'якої, ячменю звичайного, рису посівного, вівсюга звичайного та егілопса Тауша. Це стає можливим завдяки значній гомології між нуклеотидними послідовностями ділянки гена цих видів, а також тому, що РАМ-ділянка (триплет CGG) для них є однаковою. Для генетичного редагування гена *als* в ділянці 1729–1749 нуклеотидів можна створити універсальну конструкцію для усіх 7 проаналізованих видів:

пшениці м'якої, ячменю звичайного, рису посівного, вівсюга звичайного, егілопса Тауша, кукурудзи та сорго алепського. Однак, зважаючи на різний ступінь гомології, можна припустити що розроблена генетична конструкція буде працювати з різною ефективністю. Найточнішим редагування даної ділянки буде для тих генотипів, які мають між собою найвищий ступінь гомології.

### Подяки

Роботу було виконано в рамках фундаментальної відомчої тематики відділу генетичної інженерії ІКБГІ НАН України: III-8-17 «Вивчення функціонування гетерологічних генів та їх впливу на адаптаційні характеристики рослинних систем в умовах біотичних та абіотичних стресів» (№ державної реєстрації 0117U002589, 2017–2021 рр).

### Використана література

- Bo Bo A., Won O. J., Sin H. T. et al. Mechanisms of herbicide resistance in weeds. *Korean J. Environ. Agric. Sci.* 2017. Vol. 44. P. 1–15. doi: 10.7744/kjoas.2017000
- Mallory-Smith C. A., Retzinger E. J. Jr. Revised classification of herbicides by site of action for weed resistance management strategies weed technology. *Weed Technol.* 2003. Vol. 17, Iss. 3. P. 605–619. doi: 10.1614/0890-037X(2003)017[0605:RCONHS]2.0.CO;2
- Pang S. S., Duggleby R. G., Guddat L. W. Crystal structure of yeast acetoxyacid synthase: a target for herbicidal inhibitors. *J. Mol. Biol.* 2002. Vol. 317, Iss. 2. P. 249–262. doi: 10.1006/jmbi.2001.5419

4. Tan S., Evans R. R., Dahmer M. L. et al. Imidazolinone-tolerant crops: history, current status and future. *Pest Manag. Sci.* 2005. Vol. 61, Iss. 3. P. 246–257. doi: 10.1002/ps.993
5. Duggleby R. G., Pang S. S., Yu H., Guddat L. W. Systematic characterization of mutations in yeast acetohydroxyacid synthase. *Eur. J. Biochem.* 2003. Vol. 270, Iss. 13. P. 2895–2904. doi: 10.1046/j.1432-1033.2003.03671.x
6. McCourt J. A., Duggleby R. G. Acetohydroxyacid synthase and its role in the biosynthetic pathway for branched-chain amino acids. *Amino Acids.* 2006. Vol. 31, Iss. 2. P. 173–210. doi: 10.1007/s00726-005-0297-3
7. Lee H., Rustgia S., Kumara N. et al. Single nucleotide mutation in the barley acetohydroxy acid synthase (AHAS) gene confers resistance to imidazolinone herbicides. *PNAS.* 2011. Vol. 108, Iss. 21. P. 8909–8913. doi: 10.1073/pnas.1105612108.
8. Yu Q., Powles S. B. Resistance to AHAS inhibitor herbicides: current understanding. *Pest Manag. Sci.* 2014. Vol. 70, Iss. 9. P. 1–11. doi: 10.1002/ps.3710
9. Domínguez-Mendez R., Alcántara-de la Cruz R., Rojano-Delgado A. M. et al. Multiple mechanisms are involved in new imazamox-resistant varieties of durum and soft wheat. *Sci. Rep.* 2017. Vol. 7, Iss. 1. 14839. doi: 10.1038/s41598-017-13874-3
10. Pozniak C. J., Hucl P. J. Genetic analysis of imidazolinone resistance in mutation-derived lines of common wheat. *Crop Sci.* 2004. Vol. 44, Iss. 1. P. 23–30. doi: 10.2135/cropsci2004.0023
11. Anderson J. A., Matthiesen L., Hegstad J. Resistance to an imidazolinone herbicide is conferred by a gene on chromosome 6DL in the wheat line cv. 9804. *Weed Sci.* 2004. Vol. 52, Iss. 1. P. 83–90. doi: 10.1614/ws-03-055r
12. Patent US2004/0237134A1 Wheat plants having increased resistance to imidazolinone herbicides / Pozniak C. J., Hucl P. J. ; Sutherland Asbill & Brennan Llp, 999 Achtree Street, N.E. Atlanta, GA 30309 (US). Nov. 25, 2004.
13. Newhouse K. E., Smith W. A., Starrett M. A. et al. Tolerance to imidazolinone herbicides in wheat. *Plant Physiol.* 1992. Vol. 100, Iss. 2. P. 882–886. doi: 10.1104/pp.100.2.882
14. Dubey S., Bist R., Misra S. Sodium azide induced mutagenesis in wheat plant. *World J. Pharm. Pharm. Sci.* 2017. Vol. 6, Iss. 10. P. 294–304. doi: 10.20959/wjpps201710-10199
15. Jafri I. F., Khan A. H., Gulfishan M. Genotoxic effects of 5-bromouracil on cytomorphological characters of *Cichorium intybus* L. *Afr. J. Biotechnol.* 2011. Vol. 10, Iss. 52. P. 10595–10599. doi: 10.5897/AJB11.412
16. Jiang F., Doudna J. CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* 2017. Vol. 22, Iss. 46. P. 505–529. doi: 10.1146/annurev-biophys-062215-010822
17. Zhang R., Liu J., Chai Z., et al. Generation of herbicide tolerance traits and a new selectable marker in wheat using base editing. *Nat. Plants.* 2019. Vol. 5, Iss. 5. P. 480–485. doi: 10.1038/s41477-019-0405-0
18. Leach L. J., Belfield E. J., Jiang C. et al. Patterns of homoeologous gene expression shown by RNA sequencing in hexaploid bread wheat. *BMC Genomics.* 2014. Vol. 15, Iss. 276. P. 1–19. doi: 10.1186/1471-2164-15-276
19. Dvorak J., Luo M.-C., Yang Z.-L. et al. The structure of the *Aegilops tauschii* gene pool and the evolution of hexaploid wheat. *Theor. Appl. Genet.* 1998. Vol. 97, Iss. 4. P. 657–670. doi: 10.1007/s001220050942
20. Beckie H. J., Warwick S. I., Sauder C. A. Basis for herbicide resistance in canadian populations of wild oat (*Avena fatua*). *Weed Sci.* 2012. Vol. 60, Iss. 1. P. 10–18. doi: 10.1614/WS-D-11-00110.1
21. Piao Z., Wang W., Wei Y. et al. Characterization of an acetohydroxy acid synthase mutant conferring tolerance to imidazolinone herbicides in rice (*Oryza sativa*). *Planta*, 2018. Vol. 247, Iss. 3. P. 693–703. doi: 10.1007/s00425-017-2817-2
22. Wenefrida I., Utomo H. S., Meche M. M., Nash J. L. Inheritance of herbicide resistance in two germplasm lines of Clearfield rice (*Oryza sativa* L.). *Can. J. Plant Sci.* 2007. Vol. 87, Iss. 3. P. 659–669. doi: 10.4141/P05-086

## References

1. Bo Bo, A., Won, O. J., Sin, H. T., Lee, J. J., & Park, K. W. (2017). Mechanisms of herbicide resistance in weeds. *Korean J. Environ. Agric. Sci.*, 44, 1–15. doi: 10.1007/s001220050942
2. Mallory-Smith, C. A., & Retzinger, E. J. Jr. (2003). Revised classification of herbicides by site of action for weed resistance management strategies weed technology. *Weed Technol.*, 17(3), 605–619. doi: 10.1614/0890-037x(2003)017[0605:rcohbs]2.0.co;2
3. Pang, S. S., Duggleby, R. G., & Guddat, L. W. (2002). Crystal structure of yeast acetohydroxyacid synthase: a target for herbicidal inhibitors. *J. Mol. Biol.*, 317(2), 249–262. doi: 10.1006/jmbi.2001.5419
4. Tan, S., Evans, R. R., Dahmer, M. L., Singh, B. K., & Shaner, D. L. (2005). Imidazolinone-tolerant crops: history, current status and future. *Pest Manag. Sci.*, 61(3), 246–257. doi: 10.1002/ps.993
5. Duggleby, R. G., Pang, S. S., Yu, H., & Guddat, L. W. (2003). Systematic characterization of mutations in yeast acetohydroxyacid synthase. *Eur. J. Biochem.*, 270(13), 2895–2904. doi: 10.1046/j.1432-1033.2003.03671.x
6. McCourt, J. A., & Duggleby, R. G. (2006). Acetohydroxyacid synthase and its role in the biosynthetic pathway for branched-chain amino acids. *Amino Acids.*, 31(2), 173–210. doi: 10.1007/s00726-005-0297-3
7. Lee, H., Rustgia, S., Kumara, N., Burkea, I., Yenisha, J. P., Gilla, K. S., Wettsteina, D., & Ullricha, S. E. (2011). Single nucleotide mutation in the barley acetohydroxy acid synthase (AHAS) gene confers resistance to imidazolinone herbicides. *PNAS.*, 108(21), 8909–8913. doi: 10.1073/pnas.1105612108
8. Yu, Q., & Powles, S. B. (2014). Resistance to AHAS inhibitor herbicides: current understanding. *Pest Manag. Sci.*, 70(9), 1–11. doi: 10.1002/ps.3710
9. Domínguez-Mendez, R., Alcántara-de la Cruz, R., Rojano-Delgado, A. M., Fernández-Moreno, P. T., Aponte, R., & Prado, R. (2017). Multiple mechanisms are involved in new imazamox-resistant varieties of durum and soft wheat. *Sci. Rep.*, 7(10), 14839. doi: 10.1038/s41598-017-13874-3
10. Pozniak, C. J., & Hucl, P. J. (2004). Genetic analysis of imidazolinone resistance in mutation-derived lines of common wheat. *Crop Sci.*, 44(1), 23–30. doi: 10.2135/cropsci2004.0023
11. Anderson, J. A., Matthiesen, L., & Hegstad, J. (2004). Resistance to an imidazolinone herbicide is conferred by a gene on chromosome 6DL in the wheat line cv. 9804. *Weed Sci.*, 52(1), 83–90. doi: 10.1614/WS-03-055R
12. Pozniak, C. J., & Hucl, P. J. (2004). Patent US2004/0237134A1. Wheat plants having increased resistance to imidazolinone herbicides. Sutherland Asbill & Brennan Llp, 999 Achtree Street, N.E. Atlanta, GA 30309 (US). Nov. 25, 2004.
13. Newhouse, K. E., Smith, W. A., Starrett, M. A., Schaefer, T. J., & Singh, B. K. (1992). Tolerance to imidazolinone herbicides in wheat. *Plant Physiol.*, 100(2), 882–886. doi: 10.1104/pp.100.2.882
14. Dubey, S., Bist, R., & Misra, S. (2017). Sodium azide induced mutagenesis in wheat plant. *World J. Pharm. Res. Sci.*, 6(10), 294–304. doi: 10.20959/wjpps201710-10199
15. Jafri, I. F., Khan, A. H., & Gulfishan, M. (2011). Genotoxic effects of 5-bromouracil on cytomorphological characters of *Cichorium intybus* L. *Afr. J. Biotechnol.*, 10(52), 10595–10599. doi: 10.5897/AJB11.412
16. Jiang, F., & Doudna, J. (2017). CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.*, 22(46), 505–529. doi: 10.1146/annurev-biophys-062215-010822
17. Zhang, R., Liu, J., Chai, Z., Chen, S., Bai, Y., Zong, Y., ... Gao, C. (2019). Generation of herbicide tolerance traits and a new selectable marker in wheat using base editing. *Nat. Plants*, 5(5), 480–485. doi: 10.1038/s41477-019-0405-0
18. Leach, L. J., Belfield, E. J., Jiang, C., Brown, C., Mithani, A., & Harberd, N. P. (2014). Patterns of homoeologous gene expres-



- sion shown by RNA sequencing in hexaploid bread wheat. *BMC Genomics*, 15(276), 1–19. doi: 10.1186/1471-2164-15-276
19. Dvorak, J., Luo, M.-C., Yang, Z.-L., & Zhang, H.-B. (1998). The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. *Theor Appl Genet.*, 97(4), 657–670. doi: 10.1007/s001220050942
20. Beckie, H. J., Warwick, S. I., & Sauder, C. A. (2012). Basis for herbicide resistance in canadian populations of wild oat (*Avena fatua*). *Weed Sci.*, 60(1), 10–18. doi: 10.1614/WS-D-11-00110.1
21. Piao, Z., Wang, W., Wei, Y., Zonta, F., Wan, C., Bai, J., ... Fang, J. (2018). Characterization of an acetohydroxy acid synthase mutant conferring tolerance to imidazolinone herbicides in rice (*Oryza sativa*). *Planta*, 247(3), 693–703. doi: 10.1007/s00425-017-2817-2
22. Wenefrida, I., Utomo, H. S., Meche, M. M., & Nash, J. L. (2007). Inheritance of herbicide resistance in two germplasm lines of Clearfield rice (*Oryza sativa* L.). *Can. J. Plant Sci.*, 87(3), 659–669. doi: 10.4141/P05-086

УДК 602.7:57.085.2:633.11

**Кириенко А. В.<sup>1,2\*</sup>, Щербак Н. Л.<sup>1</sup>, Симоненко Ю. В.<sup>1,2</sup>** Биоинформатический анализ нуклеотидных последовательностей гена ацетолактат синтазы (*als*) различных представителей семейства Злаковых // *Plant Varieties Studying and Protection*. 2020. Т. 16, № 2. С. 207–216. <https://doi.org/10.21498/2518-1017.16.2.2020.209257>

<sup>1</sup>Институт клеточной биологии и генетической инженерии НАН Украины, ул. Академика Заболотного, 148б, г. Киев, 03143, Украина, \*e-mail: anastasija.kirienko@gmail.com

<sup>2</sup>Всеукраинский научный институт селекции, ул. Васильковская, 30, г. Киев, 03022, Украина

**Цель.** Провести биоинформатический анализ и сравнить целевые участки гена ацетолактат синтазы (*als*) у нескольких представителей семейства Злаковых и на основе полученных данных исследовать возможность создания унифицированной генетической конструкции для направленного изменения гена *als* с помощью системы CRISPR-Cas9. **Методы.** Сиквенсы гена *als* различных представителей семейства Злаковых были получены из базы данных NCBI: Nucleotide. Для сравнения был использован фрагмент гена *imi-2* пшеницы мягкой линии 'TealIMI11A' в двух участках сиквенса 367–390 и 1729–1749 нуклеотидов. Для оценки наличия нуклеотидных замен в рабочих сиквенсах гена *imi-2* использовали инструмент 'Sequence Viewer 3.37.0'. Дендрограмму строили с использованием инструмента "Blast Tree" с ресурса NCBI: Blast: Nucleotide. **Результаты.** Был проведен сравнительный анализ нуклеотидных последовательностей семи различных видов: пшеницы мягкой (*Triticum aestivum* L.), овсяга обыкновенного (*Avena fatua* L.), ячменя обыкновенного (*Hordeum vulgare* L.), риса посевного (*Oryza sativa* L.), кукурузы (*Zea mays* L.), сорго алеппского (*Sorghum halepense* Pers.) и эгилопса Тауша (*Aegilops tauschii* Coss.). На основе сравнительного анализа сиквенса участков гена *imi-2* для 7-ми генотипов построили филогенетическое дерево, которое показало, что исследованные виды можно разделить на два блока. В первый блок вошли кукуруза и сорго алеппское, а ко второму блоку – рис посевной, овсяга обыкновенный, ячмень обыкновенный, пшеница мягкая и эгилопс Тауша.

Определение степени гомологии между последовательностью 367–390 нуклеотида пшеницы мягкой и другими видами показала, что абсолютной была гомология с соответствующими последовательностями риса посевного, эгилопса Тауша и овсяга обыкновенного (100%). Наименьшим нуклеотидное родство оказалось для кукурузы и сорго алеппского – по 83,3%. На участке 1729–1749 нуклеотидов гена *imi-2* никакой из 6 сиквенсов не показал 100% гомологии с последовательностью пшеницы мягкой. Самой высокой она была для ячменя обыкновенного и эгилопса Тауша – 95,2%, а наименьшей для риса посевного, кукурузы и сорго алеппского – по 80,9%. **Выводы.** Проведенный анализ подтверждает значительную степень гомологии последовательности гена *als* для различных видов семейства Злаковых. Это позволяет допустить возможность создания универсальной генетической конструкции, с помощью которой можно осуществлять редактирование генома и получения растений, устойчивых к гербициду разных представителей этой семьи. Учитывая высокую степень гомологии последовательностей для таких видов как пшеница мягкая, эгилопс Тауша, ячмень обыкновенный, рис посевной и овсяга обыкновенный, можно предположить, что с наибольшей эффективностью соответствующая генетическая конструкция может быть использована для редактирования гена *als* именно этих генотипов.

**Ключевые слова:** пшеница мягкая; устойчивость к гербицидам; ген *als*; редактирование генома.

UDC 602.7:57.085.2:633.11

**Kyriienko, A. V.<sup>1,2\*</sup>, Shcherbak, N. L.<sup>1</sup>, & Symonenko, Yu. V.<sup>1,2</sup>** (2020). Bioinformatic analysis of nucleotide sequences of the acetolactate synthase (*als*) gene of different members of the Poaceae family. *Plant Varieties Studying and Protection*, 16(2), 207–216. <https://doi.org/10.21498/2518-1017.16.2.2020.209257>

<sup>1</sup>Institute of Cell Biology and Genetic Engineering, NAS of Ukraine, 148 Akademika Zabolotnoho St., Kyiv, 03143, Ukraine, \*e-mail: anastasija.kirienko@gmail.com

<sup>2</sup>Ukrainian Scientific Institute of Plant Breeding, 30 Vasylykivska St., Kyiv, 03143, Ukraine

**Purpose.** Provide bioinformatic analysis and comparison of target regions of the acetolactate synthase (*als*) gene in several members of the Poaceae family and, on the basis of the obtained data, explore the possibility of creating a unified genetic construct for *als* gene editing using the CRISPR-Cas9 system. **Methods.** The *als* gene sequences of various members of the Poaceae family were obtained from the NCBI: Nucleotide database. For comparison, a fragment of the *imi-2*

gene of wheat of the soft line 'TealIMI11A' was used in two regions of the 367–390 and 1729–1749 nucleotide sequences. The Sequence Viewer 3.37.0 tool was used to assess the presence of nucleotide substitutions in the working sequence of the *imi-2* gene. The dendrogram was built using the "Blast Tree" tool from the NCBI: Blast: Nucleotide resource. Results. A comparative analysis of the nucleotide sequences of seven different species was carried out: soft wheat (*Triticum aesti-*

*vum* L.), common wild oat (*Avena fatua* L.), barley (*Hordeum vulgare* L.), Asian rice (*Oryza sativa* L.), maize (*Zea mays* L.), aleppo grass (*Sorghum halepense* Pers.) and Tausch's goatgrass (*Aegilops tauschii* Coss.). The dendrogram is based on the gene sequence *als*, showed that all studied genotypes can be divided into two blocks: the first block included maize and aleppo grass, and the second block, a separate branch includes Asian rice and common wild oat, barley, soft wheat and Tausch's goatgrass. 367–390 nucleotide sequences of soft wheat showed the highest 100% homology to Asian rice, Tausch's goatgrass and common wild oat. The lowest homology was for maize and aleppo grass at 83.3%. Evaluation of the nucleotide sequence 1729–1749 showed no complete

homology at the 100% level. It was the highest for barley and Tausch's goatgrass – 95.2%, and the lowest for rice, maize and aleppo grass – 80.9% each. **Conclusions.** The analysis confirms a significant degree of homology of the *als* gene sequence for various species of the Poaceae family, which allows us to create a universal genetic vector. However, taking into account the high degree of sequence homology for species such as soft wheat, Tausch's goatgrass, barley, Asian rice and common wild oat, it can be assumed that the corresponding genetic vector can be used with the greatest efficiency to alter the *als* gene of these genotypes.

**Keywords:** *common wheat; herbicide resistance; gene als; genome editing.*

Надійшла / Received 19.05.2020  
Погоджено до друку / Accepted 10.06.2020