

ОГЛЯДИ

УДК: 577.2: 633.34

<https://doi.org/10.30970/vlubs.2020.82.01>

**МОЛЕКУЛЯРНА БУДОВА *E*-ГЕНІВ СОЇ ТА ЇХНІ
ФУНКЦІОНАЛЬНІ МУТАЦІЇ**

О. Охримович*, С. Чеботар, Г. Чеботар, Д. Жарікова

*Одеський національний університет імені І. І. Мечникова
Шампанський провулок, 2, Одеса 65058, Україна
e-mail: okhrymovych.o.v@ukr.net*

Розглянуто особливості молекулярної структури відомих *E*-локусів (early maturity) та їхньої участі в передачі сигналів до зацвітання рослин, залежно від чутливості генотипів сої до фотоперіоду. Ці локуси сприяють адаптації рослин до широкого діапазону природних умов за рахунок мутацій у генах і QTL, що контролюють час цвітіння. На молекулярному рівні *E*-гени суттєво різняться за структурними особливостями, походженням і функціями. Розміри ідентифікованих генів коливаються від одного екзону 525 п.н., кодуючого транскрипційний фактор (*E1*), до 14 екзонів і близько 20 т.п.н. для гена *GmG1a* (*E2*). Серед функціональних мутацій, які в більшості призводять до часткової або повної втрати функції, наявні одонуклеотидні заміни або делеції, вставки послідовностей, схожих на транспозони, які, у свою чергу, можуть призводити до заміни амінокислот у білку, до зсуву рамки зчитування, до появи передчасного стоп-кодона. Продукти *E*-генів є рецепторами сигналів, що надходять із навколишнього середовища, та беруть участь у сигнальних шляхах, які контролюють фотоперіод.

Загальний вплив і взаємодії між *E*-генами наразі у повному обсязі не вивчено, молекулярну будову досліджено лише для *E1-E4*, для них запропоновано генетичну мережу взаємодій, водночас п'ять локусів (*E6-E9* та *E11*) є лише картованими на хромосомах сої, а наявність окремого локусу *E5* досі не доведено. У восьми з 11 *E*-локусів домінуючий алель спричиняє пізні зацвітання. Відомо про плеiotропний вплив алелів *E*-генів на урожайність, висоту рослини, стресостійкість, реакцію на знижені температури. Знання алельного стану лише окремих із 11 генів є недостатнім. Необхідне комплексне розуміння функціонування генетичної мережі фотоперіодичної відповіді, тому що *E*-гени є генетичними детермінантами, на які можна орієнтуватися під час селекції та створення нових сортів із запрограмованими темпами розвитку.

Ключові слова: соя, *E*-гени, фотоперіодична чутливість, строки стиглості

Час цвітіння та час стиглості є важливими кількісними ознаками [1], що свідчать про географічну адаптацію, якість насіння й урожайність короткоденних рослин сої [9]. Рослини *Glycine max* L. можуть зростати в широкому діапазоні широт – щонайменше, від 50° північної до 35° південної широти [6]. Така широка адаптаційна здатність досягається за рахунок мутацій у генах і QTL, що контролюють час цвітіння [3].

У сої локуси, які відповідають за фотоперіодичну чутливість, позначаються як *E*-локуси (early maturity), вони детермінують строки стиглості цієї важливої сільськогосподарської культури. Наразі було охарактеризовано локуси *E1 – E11*. У всіх локусів, окрім *E6*, *E9* та *E11*, домінуючий алель відповідає за пізні зацвітання. Локуси *E1*, *E3* та *E4* беруть участь у відповіді на зміну фотоперіоду [28]. Локус *E1* контролює експресію генів *GmFT*, що контролюють час зацвітання (Flowering time) у *Glycine max* L. Зокрема, визначають

GmFT2a та *GmFT5a*, що є частиною сигнального шляху фітохром А. За умов довгого дня експресія *E1* активується, що, у свою чергу, викликає супресію генів *GmFT* та пізні зацвітання. Альтернативні варіанти генів *E3* та *E4*, що призводять до втрати генами функції, впливають на *E1*, знижуючи його експресію. Це, відповідно, призводить до більшої експресії генів *GmFT* і, як результат, до більш раннього цвітіння рослини незалежно від довжини дня (рис. 1) [28]. Проте запропонована [28] модель генної мережі жодним чином не відображає впливу інших *E*-генів (*E2*, *E6-11*) на час зацвітання сої.

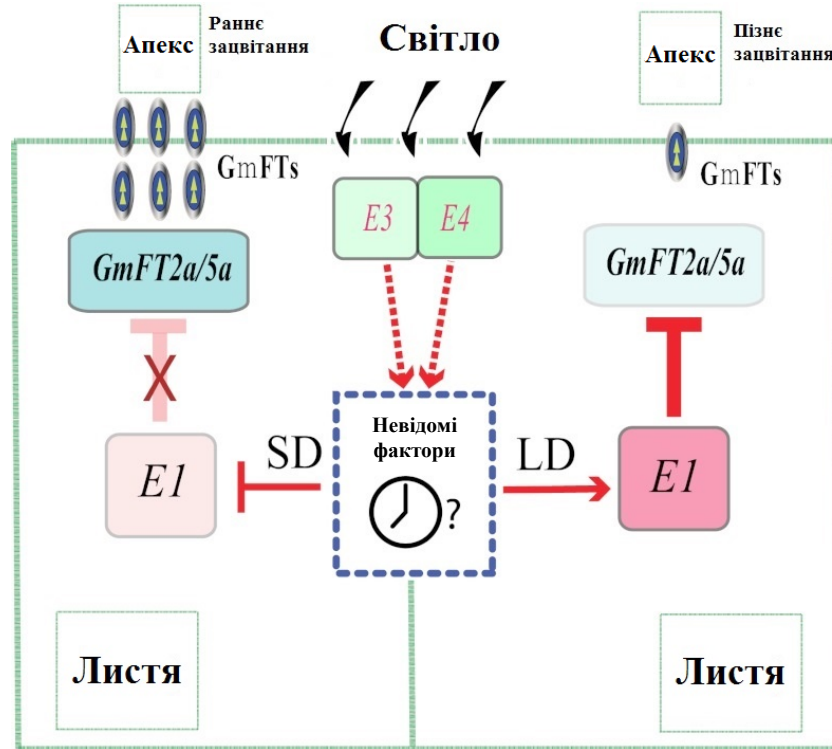


Рис. 1. Генна мережа регуляції часу цвітіння у сої [28]. Стрілками показано стимуляцію експресії генів; Т-подібні символи вказують на зниження експресії генів; X показує відміну гальмування / стимуляції; SD – коротка тривалість дня, LD – довга тривалість дня

Наша робота присвячена аналізу даних сучасної літератури щодо молекулярної структури генів *E* та їхньої участі в передачі сигналів до зацвітання рослин, залежно від чутливості генотипів сої до фотоперіоду, тому що *E*-гени є активними компонентами сигнальної системи чутливості до фотоперіоду, а білки – продукти зазначених генів – є рецепторами сигналів, що надходять із навколишнього середовища.

Молекулярна будова *E*-генів і функціональні мутації

Локус *E1* ідентифіковано за допомогою методу позиційного клонування [28] та визначено у регіоні розміром 17,4 т.п.н. на хромосомі 6 сої в перицентромерній області [21]. Цей регіон є багатим на повтори та містить мало функціональних генів [28]. Ген *E1* розміром 525 п.н. складається з одного екзону та не містить інтронів (рис. 2) [28]. *E1* кодує транскрипційний фактор, з альтернативними варіантами *el-as*, *el-fs*, *el-nl*, *el-re* та *el-p*. Для алеля *el-as* характерна наявність місенс-мутації у 44 положенні, що призводить до заміни аргініну

на треонін (*el-as*) (рис. 2). Делеція аденіну в положенні 49 алеля *el-fs* призводить до зсуву рамки зчитування й утворення передчасного стоп-кодона в положенні 124-го нуклеотиду (рис. 2). У рослин із нуль-алелем (*el-nl*) видалено ділянку 130 т.п.н., у якій міститься ген (рис. 2) [28].

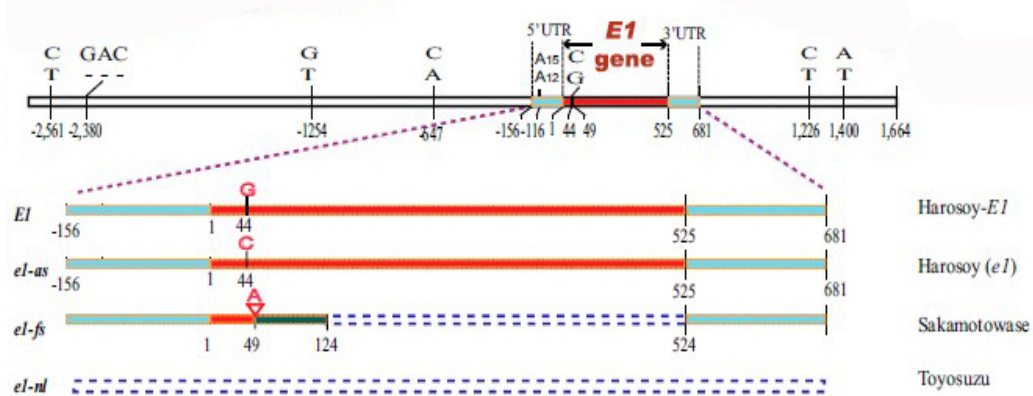


Рис. 2. Розташування та схематична будова гена *E1* згідно з результатами позиційного клонування [28]. Синій і червоний кольори позначають відповідно НТР та області кодування *E1*. Показано відмінності в молекулярній структурі алелів *E1* (клонованого з лінії Harosoy-*E1*), *el-as* (Harosoy-*el*), *el-fs* (сорту Sakamotowase), *el-nl* (сорту Toyosuzu)

В умовах Японії (Мацуда, 35°78' північної широти, 139°90' східної довготи) рослини сої з мутаціями *el-as*, *el-fs*, *el-nl* зацвітали раніше, ніж рослини дикого типу. Так, рослини-носії *el-as* зацвітали на 50-ту добу, а рослини з *el-fs* та *el-nl* на 30-ту добу після проростання, в той час як рослини з алелем дикого типу зацвітали на 70-ту добу. Якщо домінуючий алель *E1* вважається повністю функціональним алелем, алель *el-as* є частково функціональним, а алелі *el-fs* та *el-nl* є, ймовірно, нефункціональними [28]. Нещодавно було описано алелі *el-re* та *el-p*, які мають поліморфні зміни за межами генної області, що безпосередньо не впливають на структуру білка. Алель *el-re* має довгий диспергований повтор (LINE – long interspersed nuclear element) – це вставка транспозону в промоторну область *E1*, тоді як *el-p* має змінену область перед 5'-кінцем [24].

Ген *E1* кодує білок (транскрипційний фактор), який містить імовірний двосторонній сигнал ядерної локалізації та регіон, що віддалено пов'язаний зі специфічним для рослин доменом В3. Надродина В3 охоплює багато добре охарактеризованих родин з різноманітними функціями, пов'язаними з ростом і розвитком рослин [22]. Наприклад, ARF (фактори відповіді на ауксин), ABI3 (фактор нечутливості до абсцизової кислоти) у арабідопсису. Домен В3 відповідає за фактори транскрипції у більш ніж 40 видів рослин. Він складається зі 100–120 залишків, включає сім бета-ланцюгів і дві альфа-спіралі, які утворюють ДНК-зв'язуючу білкову частину [29]. Білок, що кодується *E1* геном, лише на 21–27 % ідентичний до В3 домену, знайденого в багатьох транскрипційних факторах різних видів рослин [16]. Він вважається супресором експресії генів *GmPhyA3* та *GmPhyA2*, які є ортологами локусу арабідопсису під назвою *FLOWERING TIME (FT)* [11]. Гени *GmPhyA3* та *GmPhyA2* регулюються локусами *E3* та *E4* [28]. С-кінцевий регіон амінокислотної послідовності (55–174) білка, що кодується геном *E1*, несе мінімальний В3 домен, необхідний для контакту з ДНК [22, 29]. Характеристики, виявлені за допомогою досліджень *in silico*, вказують на те, що білок, який кодується геном *E1*, може функціонувати у зв'язуванні ДНК або як фактор

транскрипції та містить два основних домени (KKRK і RRR), що розділені 12 амінокислотними залишками [28].

Було показано, що експресія гена *E1* є найбільшою у листках рослин сої, які повністю розпустилися [28]. Протягом дня експресія *E1* змінюється поступово, вночі вона зменшується і досягає мінімуму перед світанком [28].

Вважається, що гени часу цвітіння, як правило, мають плейотропну дію на такі важливі ознаки, як урожайність, висота рослини та стресостійкість. Як було відмічено вище, QTL, асоційовані з генетичними локусами, що детермінують підвищену урожайність і реакцію на знижені температури (менше 17 °C), тісно пов'язані з локусом *E1*. Однак, на думку Funatsuki et al. [8], подальші дослідження дають змогу виявити, чи є це впливом самого *E1* гена, чи ж незалежних сусідніх генів.

У локусі сої *E2* за допомогою позиційного клонування на хромосомі 10 виявлено ген *GmGla* (*Glyma10g36600*) – ортолог гена арабідопсису *GIGANTEA* [27]. *GmGla* кодує білок *GIGANTEA* (GI), що відіграє важливу роль у цвітінні завдяки контролю рівня експресії м-РНК [12, 18].

Кодуюча ділянка гена розміром 20 т.п.н. містить 14 екзонів (рис. 3,А) [27]. До алельних варіацій локусу *E2* належить нонсенс-мутація (*e2-ns* або *ft2*) – зміна А на Т в положенні 1561, яка вводить передчасний стоп-кодон (рис. 3, Б). До цього локусу розроблено алель-специфічний маркер [27]. У мутантній лінії *E2-mut* делеція нуклеотидів у 10-му екзоні призводила до трансляції усіченого білка GI. Такі рослини характеризувалися більш ранніми строками цвітіння у природних умовах [27].

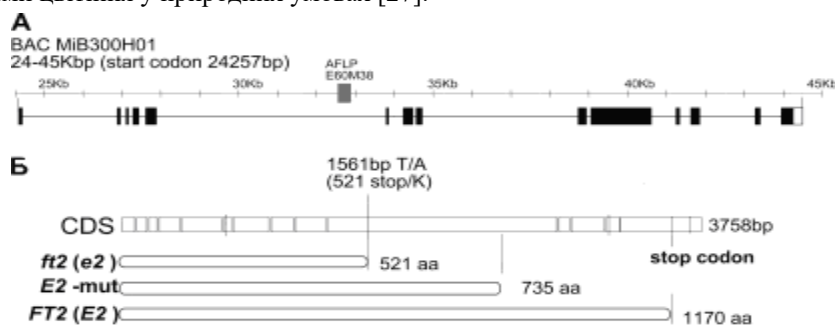


Рис. 3. Структура гена *GmGla* та положення мутацій: А – схематична структура гена *GmGla* [27]. Екзони позначені прямокутниками. Б – відмінності в молекулярній структурі алелів *ft2*, *E2-mut*, *FT2* (*E2*)

У дослідженнях [27], проведених у Японії, різниця між близько ізогенними лініями у часі цвітіння була достовірною ($P < 0,001$) з $57,5 \pm 1,72$ добу ($n=168$) для рецесивного алеля та $67,6 \pm 1,56$ ($n=210$) доби для домінантного алеля. Тобто генотип *e2/e2* спричиняв раннє цвітіння, викликаючи експресію *GmFT2a*, що є гомологом флоригену в сої. Ці результати вказують на те, що функції *GI* у цвітінні зберігаються в сої, а мутації з втратою функцій гена *GI* можуть бути корисними ресурсами для адаптації рослин зі складними геномами, такими як соя, до широкого кола географічних регіонів [27].

У локусі *E3*, який локалізовано на хромосомі 19 геному сої, методом позиційного клонування, було визначено наявність гена фітохром А сої – *GmPhyA3* [26], який містить 4 екзони. Добре відомо, що мутації гена фітохром А впливають на чутливість рослин до

фотоперіоду [26]. Білок, що кодується геном *GmPhyA3*, є фоторецептором шляху цвітіння, який затримує цвітіння в умовах довгого дня з високою інтенсивністю червоного (R, довжина хвилі 625–740 нм) / далеко-червоного світла (FR, довжина хвилі 700-780 нм) [15, 26].

У дослідженнях [26], проведених у Японії, різниця за часом цвітіння була істотною ($P < 0,001$) між близько-ізогенними лініями, що різнилися алелями *e3* та *E3*. Лінії з рецесивним алелем *e3* зацвітали на $58,2 \pm 1,55$ (n=47) добу, а лінії з домінантним алелем *E3* – на $69,1 \pm 1,21$ добу (n=70).

Одна з алельних форм (*GmPhyA3-Mo*) локуса *E3* містить однунулеотидну несинонімічну амінокислотну заміну (гліцин на аргінін) у третьому екзоні та велику інсерцію у четвертому інтроні. Послідовність інсерції має довжину 2,5 т.п.н., а її частина дуже схожа на довгу кінцеву послідовність ретротранспозону, що кодує зворотну транскриптазу. Однак функціональне значення цієї вставки залишається незрозумілим [26]. У іншій алельній формі (*GmPhyA3-E3*) міститься велика інсерція, що теж схожа на послідовність ретротранспозонів. Рослини з цим алелем зацвітали на 15 днів раніше, порівняно з рослинами з алелем дикого типу (в експерименті рослини вирощували під штучним освітленням ртутною лампою). У третього алельного варіанта (*GmPhyA3-e3*) було детектовано велику делецію (13,33 т.п.н.) після третього екзона. Можливо, це призводить до експресії нефункціонального фітохрому, оскільки в нього було видалено гістидинкіназний домен, що відіграє важливу роль у передачі сигналу [26]. Структура гена *GmPhyA3* відображена на рис. 4.

У рису подвійні мутанти *phyA-phyB* та *phyA-phyC*, вирощені в природних умовах, проявили фенотип більш раннього зацвітання, а ефект *phyA* маскувався за наявності інших функціональних фоторецепторів. Ця епістатична взаємодія, що спостерігається у рису, може бути порівнянна з епістазом між локусами *FT1 (E1)* та *FT3 (E3)* у сої [26]. Більш того, ймовірно, що *GmPhyA3* ген набув здатності сприймати червоне світло в результаті субфункціоналізації після подвоєння геному в сої. Однак доказів цього немає, і для перевірки гіпотези необхідно більше експериментів [26].

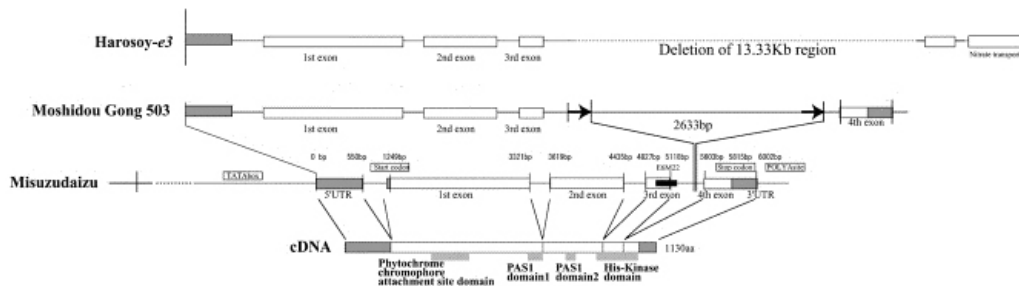


Рис. 4. Структура гена *GmPhyA3* згідно з [26]. Білі прямокутники, сірі прямокутники та горизонтальні лінії позначають відповідно екзони, НТР та інтрони; вказно позиції ТАТА-подібної послідовності, ініціаторного кодона та стоп-кодона. Видалену область, виявлену в *Harosoy-e3*, позначено пунктирною лінією. Цифри над алелем сорту *Misuzudaizu* вказують на нуклеотидну позицію від початкової точки транскрипції. Положення великої інсерції, що спостерігається в генотипі *Moshidou Gong 503*, позначене жирними стрілками, що вказують праворуч

У даний час інформація про фізіологічну активність білків фітохрому А сої та їхній взаємозв'язок з іншими продуктами генетичного годинника, а також про взаємодію з іншими білками сої, обмежена. Певною мірою високий рівень подібності між повторюваними

генами може перешкоджати аналізу функцій окремих генів *GmPhyA3*. Для з'ясування контролю за цвітінням сої необхідно виділити різні гени для асоційованих локусів і проаналізувати їхні функції та взаємодії [26].

У локусі *E4*, який міститься на 20-й хромосомі геному сої, наявний інший ген фітохрому А (*GmFT5a*), що містить 5 екзонів [15]. Він має 1 функціональний і 5 нефункціональних алелів [13], до яких розроблені алель-специфічні праймери [15]. У рецесивному алелі (*e4-SORE-1*) наявна інсерція 6238 п.н. – ця ділянка схожа на ретротранспозон (*Ty1/copia*) [22, 23] (рис. 5). Її наявність вносить передчасний стоп-кодон, який вкорочує 1123 амінокислотний протеїн до 237 амінокислот. Це порушує функцію *GmFT5a* гена щодо забезпечення нечутливості до довгих світлових днів із низьким співвідношенням R: FR [15]. Інших чотири дисфункціональних алелі – *e4-oto*, *e4-tsu*, *e4-kam* і *e4-kes* – мають по одному видаленому нуклеотиду в гено-кодуєчій області, що створює зміщення рамки зчитування та експресію усічених білків з 456 амінокислот, 759 амінокислот, 894 амінокислот і 979 амінокислот відповідно [15].

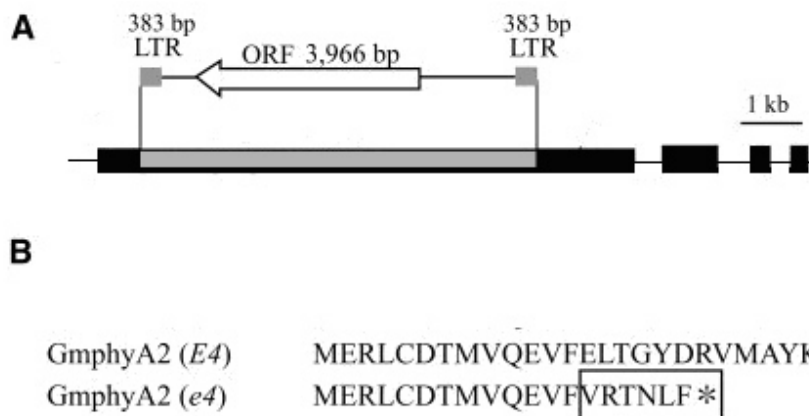


Рис. 5. Структура *GmphyA2* [15]: А – Суцільні чорні та сіре поля вказують на екзони і ретротранспозон *Ty1/copia*, відповідно. Біла стрілка вказує положення та напрямок відкритої рамки зчитування (ORF) в ретротранспозоні. В – Зміни у вирівняній послідовності амінокислот, які були викликані введенням ретротранспозону в *GmphyA2* для алеля *e4* порівняно з *E4*. Послідовність амінокислот ретротранспозона позначена квадратом. Зірочка вказує стоп-кодон [15]

Зв'язок між локусами *E3* та *E4* у сої може бути еквівалентним взаємодії *phyA* та *phyB* або *phyC* у рису [15]. Разом з алелем *e3* алель *e4*, швидше за все, обумовлює нечутливість до фотоперіоду в сої [15]. Скоріш за все, *GmFT2a* та *GmFT5a* – це гомеологи, що виникли внаслідок давніх хромосомних дублювань і перебудов у геномі сої [15].

Локус *E5* – на сьогодні немає чітко визначеної думки щодо підтвердженого його існування. Дослідники [17] припускали, що *E5* міститься на 10-й хромосомі геному сої у безпосередній близькості від гена *E2*, або ж *E5* є алелем локусу *E2*. *E5* показав часткове домінування, аналогічне повідомленому для алелів *E1* і *E2* [7], а також відрізнявся за своїм впливом від *E2* та *E1* на час цвітіння та стиглості [7].

У той же час [17] після дослідження близько-ізогенних ліній, створених на основі сортів Harosoy, Clark, а також японської лінії PI 80837, зробили висновок, що унікального гена *E5* немає, хоча існування нового локусу, який впливає на строки стиглості сої, не може бути виключеним.

Локус E6. Повідомлялося, що є ген, який відіграє важливу роль у подовженні ювенільної фази (Long juvenile) та затримує цвітіння в умовах короткого дня. Ознаку було виявлено у тропічних сортах сої, а подальше її введення в інші сорти зробило вирощування сої можливим навіть у регіонах із широтами нижче 15° південної широти [2].

Уперше дослідження нового локусу сої E6 було проведено в експериментах Cober, Voldeng [5], але остаточних доказів його існування та наявності алельних форм не зроблено. Комітет із генетики сої переглянув представлену інформацію та призначив символ E6 локусу, що визначає раннє цвітіння та стиглість, характерну для сорту Paraná, а e6 – гена, який визначає пізнє цвітіння і стиглість у сорту Paranagoiana та лінії SS-1 [2].

У подальших дослідженнях [2] доведено, що цей ген міститься на хромосомі 4 поряд із маркером HRM101. Також було показано, що гени E6 та E1 мають епістатичну взаємодію, причому E6 є супресором гена E1.

Локус E7 визначає чутливість сої до фотоперіоду і є тісно зчепленим з E1. E7 міститься на хромосомі 6 геному сої, на відстані 6,0 сМ від локусу E1 [5]. Генотип E7/E7 призводить до більш пізнього цвітіння та стиглості, чутливості до тривалих фотоперіодів, тоді як e7/e7 призводить до раннього цвітіння та стиглості, меншої чутливості до довгого фотоперіоду [5]. Локус E7 містить діагностичний маркер Satt319 та займає ділянку 22,2 сМ між маркерами Satt100 і Satt469 [19]. На сьогодні даних щодо секвенування нуклеотидної послідовності цього гена та визначення його молекулярної структури немає, як і відомостей про будову та функціонування продукту цього гена, тому поки що неможливо визначити місце цього гена в генній мережі регуляції чутливості рослин сої до фотоперіоду.

Локус E8 міститься в прецентромній області хромосоми 4 геному сої. Він специфічно відображається у групі зчеплення C1 між маркерами Sat 404 та Satt136. Генотип E8/E8 призводить до більш пізньої стиглості, тоді як e8/e8 – до дострокової стиглості [4]. Рослини сої з генотипами e8/e8 дозрівали на 9 і 6 днів раніше, що показано для сортів Maple Presto та Harosoy відповідно, за вирощування в умовах Оттави (Канада) [4]. На обох генетичних фонах лінії з рецесивним алелем e8 характеризувалися меншою кількістю насіння, давали близько 60 % врожаю та мали висоту рослин на рівні 80 % висоти ліній із домінантним алелем [4]. Подальші дослідження зі секвенування цільової ділянки та пошуку гена E8 і визначення його молекулярної будови дадуть змогу більш повно зрозуміти його роль у розвитку рослин сої.

Локус E9 міститься на хромосомі 16 геному сої в регіоні довжиною ~245 т.п.н. Є два гени-претенденти, що можуть міститися в цьому локусі: GmFT2a та GmFT2b. Для точного визначення, який із цих генів є у локусі E9, необхідні подальші дослідження [10]. Дослідники [10], використовуючи лінії Hidaka 4, Tokei 780 та ранньоквітучі рекомбінантні інбредні лінії (RIL) TH85 та TH03, вирощені в умовах Китаю (Харбін, 45° 43' Північної широти, 126° 45' Східної довготи), визначили, що домінантний алель E9/E9 обумовлює раннє цвітіння та стиглість за умов довгого дня (LD).

Локус E10 кодує T-подібний білок 8. Аналіз простих повторюваних послідовностей (SSR) та однонуклеотидних поліморфізмів (SNP) визначив область на хромосомі 8 (група зчеплення A2), що корелює з цим новим локусом стиглості. Використовуючи функціональну геноміку сої (за допомогою аналізу PIPE) з подальшим застосуванням методів молекулярної біології, які включали секвенування, аналіз SNP та 2D дослідження РНК, було визначено, що FT4 є геном-кандидатом для локусу скоростиглості E10 [20]. Генотип E10/E10 призводить до пізньої, а e10/e10 – до ранньої стиглості. Скоріш за все, заміщення треоніну ізолейцином у 4-му екзоні призводить до зміни функції білка рослин із мутантним алелем [20].

Локус *E11* на хромосомі 7 також відіграє важливу роль у регуляції часу цвітіння та дозрівання у сої (рис. 6).

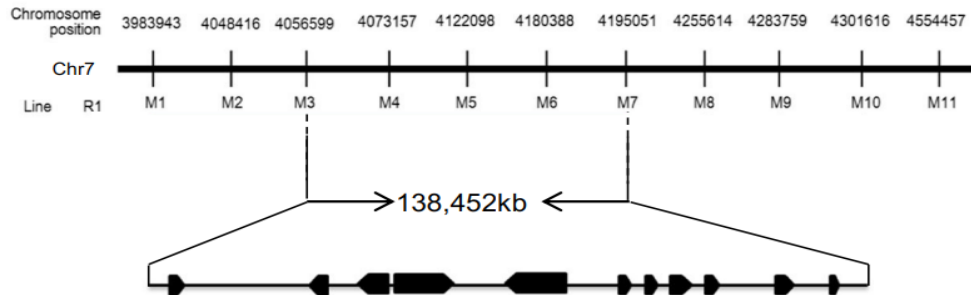


Рис. 6. Приблизне положення гена *E11* на хромосомі 7 у області довжиною близько 138 т.п.н. між маркерами M3 та M7. Жирними чорними стрілками вказано відкриті рамки читування в цільовому регіоні [25]

Дослідники [25] вважають, що цей локус може міститися між індел-маркерами (indel markers) *ID7079* та *ID7088* у області довжиною 138 т.п.н., в якій запропоновано гені-кандидати *Glyma.07 g4850*, *Glyma.07 g049000* чи *Glyma.07 g049100* [25].

Локус *E11* має подібну до локусу *E9* дію, його домінуючий алель викликає раннє зацвітання в умовах довгого дня [25].

Узагальнення

Таким чином, у роботі розглянуто й узагальнено особливості молекулярної структури відомих *E*-генів сої. Функціональні особливості зазначених генів впливають на час цвітіння рослин сої, що свідчить на користь їхньої участі у фотоперіод-сигнальних шляхах рослин. *E*-гени є генетичними детермінантами, на які можна орієнтуватися під час селекції та створення нових сортів із запрограмованими темпами розвитку. Водночас стає зрозумілим, що знання алельного стану лише окремих з 11 генів недостатньо. Необхідне комплексне розуміння функціонування генетичної мережі фотоперіодичної відповіді. Нещодавно опубліковані факти дають змогу створити часткову модель взаємодії досліджуваних генів, як, наприклад, у роботі Xia et al. [28]. За потрапляння світла на листову пластинку сигнал передається завдяки продуктам генів фітохромів *E3* та *E4*, проте в умовах короткого дня не відбувається активації експресії *E1*, що знімає репресію генів *GmFT*, які контролюють час зацвітання (Flowering time) у *Glycine max* L., та призводить до більш раннього цвітіння [28]. Продукти мутантних генів *E3* та *E4* через втрату функціональності призводять до змін на шляху передачі світлового сигналу, в результаті знижується рівень експресії транскрипційних факторів (*E1*) – супресорів *GmFT*, що призводить до більшої експресії генів *GmFT* і, як результат, до більш раннього цвітіння рослини незалежно від довжини світлового дня.

За вирощування рослин сої в Україні важливим для селекціонерів є створення саме ранньостиглих сортів, що дає змогу більш інтенсивно використовувати сою як проміжну та повторну культуру.

Залишаються незрозумілими вплив інших *E*-генів та їхня участь у формуванні чутливості до фотоперіоду і скорочення часу до цвітіння. Гени *E6-E9* та *E11* на сьогодні тільки картовані на хромосомах сої, тоді як існування *E5* досі викликає сумніви. Беззаперечно складність кодування часу і тривалості цвітіння й довжини вегетаційного періоду потребує від дослідників подальших робіт зі секвенування генів і визначення їхньої молекулярної структури та формування на основі отриманих даних прототипів генних мереж взаємодії *E*-генів один з одним або з іншими генами.

Залишається багато невирішених питань. Так, на сьогодні є інформація про молекулярну будову лише генів *E1-E4*. Для *E7*, наприклад, відома приблизна локалізація на хромосомі, проте наявність мутацій або білкова структура залишаються невідомими, така ж ситуація спостерігається і для *E8*.

У той же час в дослідженнях українських зразків сої [30, 30] за допомогою маркерів до мікросателітних локусів, зчеплених з *E7*, простежується тенденція до більш частого трапляння саме домінантних алелів гена *E7* в українському генофонді, а не інших *E*-генів.

Отже, на сьогодні ще залишається багато питань, які стосуються вдосконалення наших уявлень про генетичні мережі, що контролюють час цвітіння у сої.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Bernard R. L. Two major genes for time of flowering and maturity in soybeans // *Crop Science*. 1971. Vol. 11. P. 242–247.
2. Bonato E. R., Vello N. A. *E6*, a dominant gene conditioning early flowering and maturity in soybeans // *Genet. Mol. Biol.* 1999. Vol. 22. P. 229–232.
3. Cao D., Takeshima R., Zhao C. et al. Molecular mechanisms of flowering under long days and stem growth habit in soybean // *J. Exp. Bot.* 2017. Vol. 68. N 8. P. 1873–1884
4. Cober E. R., Molnar S. J., Charette M., Voldeng H. D. A new locus for early maturity in soybean // *Crop Sci.* 2010. Vol. 50. P. 524–527.
5. Cober E. R., Voldeng H. D. A new soybean maturity and photoperiod-sensitivity locus linked to *E1*, and *T* // *Crop Sci.* 2001. Vol. 41. N 3. P. 698–701.
6. Destro D., Carpentieri-Pipolo V., Kiihl R. A. S., Almeida L. A. Photoperiodism and genetic control of the long juvenile period in soybean: A review // *Crop Breed. Appl. Biotechnol.* 2001. Vol. 1. P. 72–92.
7. Dissanayaka A., Rodriguez T. O., Di S. Quantitative trait locus mapping of soybean maturity gene *E5* // *Breed Sci.* 2016. Vol. 66. N 3. P. 407–415.
8. Funatsuki H., Kawaguchi K., Matsuba S., Sato Y., Ishimoto M. Mapping of QTL associated with chilling tolerance during reproductive growth in soybean // *Theor. Appl. Genet.* 2005. Vol. 111. P. 851–861.
9. Garner W. W., Allard H. A. Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants // *J. Agric. Res.* 1920. Vol. 18. P. 553–606.
10. Kong F., Nan H., Cao D. et al. New Dominant gene *E9* conditions early flowering and maturity in soybean // *Crop Science*. 2014. Vol. 54. P. 2529–2535.
11. Kong F.J., Liu B., Xia Z. Two coordinately regulated homologs of *FLOWERING LOCUS T* are involved in the control of photoperiodic flowering in soybean // *Plant Physiol.* 2010. Vol. 154. P. 1220–1231.
12. Koornneef M., Alonso-Blanco C., Blankestijn-de Vries H. et al. Genetic interactions among late-flowering mutants of *Arabidopsis* // *Genetics*. 1998. Vol. 148. N 2. P. 885–892.
13. Langewisch T., Zhang H., Vincent R. et al. Major soybean maturity gene haplotypes revealed by SNPviz analysis of 72 sequenced soybean genomes // *PLOS ONE*. 2014. Vol. 9. N 4. P. e94150.
14. Li X., Fang C., Xu M. et al. Quantitative trait locus mapping of soybean maturity gene *E6* // *Crop Science*. 2017. Vol. 57. N 5. P. 1–8.
15. Liu B., Kanazawa A., Matsumura H. et al. Genetic redundancy in soybean photoresponses associated with duplication of phytochrome *A* gene // *Genetics*. 2008. Vol. 180. P. 995–1007.

16. Lukashin A.V., Borodovsky M. GeneMark, new solutions for gene finding // *Nucleic Acids Res.* 1998. Vol. 26. N 4. P. 1107–1115.
17. McBlain B. A., Bernard R. L. Anewgene affecting the time of flowering maturity in soybeans // *J. Hered.* 1987. Vol. 178. P. 68–70.
18. Mizoguchi T., Wright L., Fujiwara S. et al. Distinct roles of *GIGANTEA* in promoting flowering and regulating circadian rhythms in Arabidopsis // *Plant Cell.* 2005. Vol. 17. N 8. P. 2255–2270.
19. Molnar S. J., Rai S., Charette M., Cober E. R. Simple sequence repeat (SSR) markers linked to *E1*, *E3*, *E4*, and *E7* maturity genes in soybean // *Genome.* 2003. Vol. 46. P. 1010–1024.
20. Samanfar B., Molnar S. J., Charette M. et al. Mapping and identification of a potential candidate gene for a novel maturity locus, *E10*, in soybean // *Theor. Appl. Genet.* 2017. Vol. 130. N 2. P. 377–390.
21. Schmutz J. Genome sequence of the palaeopolyploid soybean // *Nature.* 2010. Vol. 463. P. 178–183.
22. Swaminathan K., Peterson K., Jack T. The plant B3 superfamily // *Trends Plant Sci.* 2008. Vol. 13. N 12. P. 647–55.
23. Tsubokura Y., Matsumura H., Xu M. et al. Genetic variation in soybean at maturity locus *E4* is involved in adaptation to long days at high latitudes // *Agronomy.* 2013. Vol. 3. P. 117–134.
24. Tsubokura Y., Watanabe S., Xia Z. et al. Natural variation in the genes responsible for maturity loci *E1*, *E2*, *E3* and *E4* in soybean // *Ann. Bot.* 2014. Vol. 113. N 3. P. 429–441.
25. Wang F., Nan H., Chen L. et al. A new dominant locus, *E11*, controls early flowering time and maturity in soybean // *Mol. Breeding.* 2019. Vol. 39. P. 70.
26. Watanabe S., Hideshima R., Xia Z. Map-based cloning of the gene associated with the soybean maturity locus *E3* // *Genetics.* 2009. Vol. 182. P. 1251–1262.
27. Watanabe S., Xia Z., Hideshima R. A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering // *Genetics.* 2011. Vol. 188. P. 395–407.
28. Xia Z., Watanabe S., Yamada T. Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering // *Proc. Nat. Acad. Sci., USA.* 2012. Vol. 109. P. E2155–E2164.
29. Yamasaki K., Kigawa T., Inoue M. et al. Solution structure of the B3 DNA binding domain of the Arabidopsis cold-responsive transcription factor RAV1 // *Plant Cell.* 2004. Vol. 16. N 12. P. 3448–3459.
30. Zharikova D. O., Chebotar G. O., Aksonova O. A., Chebotar S. V. Vyznachennia alelnoho stanu heniv *E1*, *E2*, *E3*, *E4* ta *E7* za dopomohoiu zcheplenykh z nymy mikrosatelitnykh lokusiv u sortiv soi suchasnoi selektsii // *Ekoloho-henetychni aspekty v selektsii polovykh kultur v umovakh zmin klimatu: Materialy mizhnarodnoi naukovy-praktychnoi konferentsii, prysviachenoi 90-richchiu z dnia narodzhennia henetyka, selektsionera, profesora M.M. Chekalina (18–19 kvitnia 2019 r.) / Poltavska derzhavna ahrarna akademiia. Poltava, 2019. P. 29.*
31. Zharikova D., Chebotar G., Aksonova E. et al. Polymorphism in SSR-loci associated with *E* genes in soybean mutant lines perspective for breeding // *Agricultural science and practice.* 2019. Vol. 6 (3). P. 45–55.

Стаття надійшла до редакції 09.01.20

прийнята до друку 18.02.20

**MOLECULAR STRUCTURE OF SOYBEAN *E*-GENES AND THEIR
FUNCTIONAL MUTATIONS**

O. Okhrymovych, S. Chebotar, G. Chebotar, D. Zharikova

*Odesa National I. I. Mechnikov University
Shampansky Lane, 2, Odesa 65058, Ukraine
e-mail: okhrymovych.o.v@ukr.net*

In this review, we discuss features of the molecular structure of known *E*-loci (early maturity) and their involvement in signaling to plant flowering, depending on the sensitivity of soybean genotypes to the photoperiod. These loci contribute to the adaptation of plants to a wide range of natural conditions due to mutations in genes and QTL that control flowering time. At the molecular level, *E*-genes are significantly different in structural features, origin and function. The length of the identified genes range from one exon to 525 bp encoding the transcription factor (*E1*), up to 14 exons and about 20 kb for the *GmGla* gene (*E2*). Among the functional mutations that in most cases lead to partial or complete loss of function, there are single-nucleotide substitutions or deletions, insertions of transposon-like sequences that can lead to amino acid substitutions in the protein, shift of the reading frame, appearance of the premature stop-codon. *E*-gene products are receptors of signals coming from the environment and they participate in signaling pathways that control the photoperiod.

The overall impact and interactions between *E*-genes have not been fully studied yet, the molecular structure was investigated only for *E1-E4*, for which a genetic network of interactions was proposed, while at the same time five loci (*E6-E9* and *E11*) were only mapped on soybean chromosomes, and the existence of a separate *E5* locus has not yet been established. In eight of the 11 *E*-loci, the dominant allele causes late flowering. Also there is a pleiotropic effect of *E*-gene alleles on yield, plant height, stress resistance, and response to low temperatures. Knowledge of the allelic state of only some of the 11 genes is not sufficient. A comprehensive understanding of the functioning of the photoperiodic genetic response network is needed. *E*-genes are genetic determinants that can be used during selection and creation of new varieties with programmed rates of development.

Keywords: soybean, *E*-genes, photoperiodic sensitivity, maturity