

АНОТАЦІЯ

Варченко О.І. Вивчення гетерологічної експресії репортерного гена *gfp* в рослинах *Nicotiana rustica* L. – Кваліфікаційна наукова праця на правах рукопису. Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 091 «Біологія» (09 «Біологія») – Інститут клітинної біології та генетичної інженерії НАН України, Київ, 2021.

Сучасні напрямки досліджень вітчизняних та закордонних учених демонструють, що для генетичної інженерії рослин важливим завданням є поліпшення їх господарсько-цінних ознак. Успішне виконання такого завдання пов'язане з ефективною та контрольованою експресією перенесених генів. Експресія генів, в свою чергу, залежить від вибору видів-хазяїв, регуляторних елементів, що забезпечують функціонування гетерологічних генів, феномену «сайленсингу» генів. Останнім часом, дослідники використовують не тільки стабільну генетичну трансформацію, але й транз'єнтну. При транз'єнтній трансформації не відбувається інтеграції перенесених генів в ядерну ДНК, а експресія гетерологічних генів проходить тимчасово. Цей метод дає змогу швидко (протягом декількох днів) аналізувати рівень експресії гетерологічних генів та визначати придатність нових генетичних конструкцій для отримання трансгенних рослин з поліпшеними ознаками. Використовуючи та порівнюючи стабільну і транз'єнтну експресію генів, вчені отримали унікальну можливість досліджувати фундаментальні питання молекулярної біології та прикладної генетичної інженерії. Однак, наразі в генно-інженерних дослідженнях використовується мала різноманітність регуляторних елементів, що не забезпечує дослідників потрібними знаннями про діапазон експресійних можливостей генів з використанням тих чи інших елементів генетичної конструкції.

Дисертаційне дослідження присвячене вивченню особливостей експресії перенесених генів в рослинах, впливу регуляторних елементів різного походження (рослинних, вірусних, бактеріальних) на рівень експресії гетерологічного гена *gfp*, що кодує зелений флуоресцентний білок

(*green fluorescent protein*), систематизації даних щодо рівня експресії гетерологічного гена *gfp*, використовуючи транз'єнтну трансформацію махорки *Nicotiana rustica* L. (*N. rustica*).

У дисертаційній роботі використано такі методи досліджень: молекулярне клонування Golden Gate; рестрикційний аналіз ДНК, полімеразно ланцюгова реакція; приготування компетентних бактеріальних клітин *Escherichia coli* L. (*E. coli*) та *Agrobacterium tumefaciens* L. (*A. tumefaciens*); трансформація компетентних клітин *E. coli* та *A. tumefaciens*; транз'єнтна трансформація рослин; спектрофлуориметричний аналіз для визначення рівня накопичення флуоресцентного рекомбінантного білка GFP; кількісний аналіз білків за методом Бредфорда; електрофоретичне розділення білків у поліакриламідному гелі; денситометричний аналіз; комп'ютерне моделювання *in silico*, метод математичної статистики (для оцінки достовірності одержаних результатів).

Аналіз літературних даних свідчить про можливість використання рослинних систем для синтезу рекомбінантних білків, які можуть слугувати моделлю для дослідження генів та їх регуляторних елементів, що сприяє поглибленню наших знань щодо механізмів регуляції експресії генів. Останнім часом було перевірено ефективність використання різних видів сільськогосподарських культур (рис, пшениця, люцерна, картопля) як об'єктів для транз'єнтної експресії генів, але рослини роду *Nicotiana* L. (тютюни) виявились найефективнішими для агробактеріальної транз'єнтної трансформації. Серед видів цього роду найчастіше для досліджень використовують модельний вид *Nicotiana benthamiana* L. Крім цього, в експериментах використовують інші види, такі як *N. excelsior*, *N. tabacum*, гібрид *N. excelsiana*. Для порівняння рівнів експресії гетерологічних генів при тестуванні генетичних конструкцій потрібні високі рівні накопичення рекомбінантних білків в рослинних системах, стабільність цільових білків в рослинних клітинах, зручність при вирощуванні видів-хазяїв та проведенні експериментів. Рослини виду *N. benthamiana*, зазвичай, демонструють високий рівень експресії цільового гена, але мають невеликий розмір листових пластин. Потенційні розміри та біомаса листків махорки *N. rustica* суттєво перевищують

N. benthamiana, а рівень експресії гетерологічного гена *gfp* співставний, згідно з літературними даними та отриманими результатами досліджень. Використовуючи для інфільтрації лише одну листову пластину рослин *N. rustica*, можливо проаналізувати 12-15 конструкцій одночасно, що в тричотири рази вище порівняно із використанням *N. benthamiana*. Інші переваги використання махорки полягають у її невибагливості та легкості вирощування в теплиці, тому як основний об'єкт дослідження нами було обрано саме вид *N. rustica*.

Для дослідження гетерологічної експресії гена *gfp* в рослинах виду *Nicotiana rustica* L. нами було створено 19 генетичних конструкцій. Використовували наступні генетичні конструкції для визначення впливу різних регуляторних елементів на рівні експресії гена *gfp*: вектори pSPV2301-pSPV2306 для перевірки промоторів; вектори pSPV2303, pSPV2307-pSPV2309 для перевірки сигнальних пептидів, вектори pSPV2303, pSPV2310-pSPV2312, pSPV2317, pSPV2318 для перевірки 5'-последовностей, що не транлюються; вектори pSPV2312-pSPV2316 для перевірки термінаторів; pSPV2313 та pSPV2324 для перевірки впливу білка-супресора посттранскрипційного сайленсингу генів.

В роботі визначено оптимальну концентрацію агробактеріальної суспензії для транз'єнтної трансформації рослин *N. rustica*. Для цього використовували генетичну конструкцію pSPV2303, що несе ген *gfp* під контролем подвійного 35S промотора вірусу мозаїки цвітної капусти (*Cauliflower Mosaic Virus*) (D35SCaMV P), 5'-последовності омега, що не транлюється вірусу тютюнової мозаїки (*Tobacco Mosaic Virus*) (5'Ω TMV) та 35S термінатора вірусу мозаїки цвітної капусти (35S T). Для досліджень використовували бактеріальну суспензію з оптичною густиною 0,2; 0,4; 0,8; 1,0; 1,5; 2,0 при довжині хвилі 600 нм (OD₆₀₀) у варіантах без та з додаванням ацетосирингону (індуктора бактеріальних генів вірулентності) у концентрації 150 мкмоль. Визначили, що використання ацетосирингону в концентрації 150 мкмоль статистично не впливало на рівень експресії гена *gfp*, що підтверджується спектрофлуориметричним аналізом рівнів флуоресценції білка GFP та

визначенням концентрації загальних водорозчинних білків методом Бредфорда. Тому детекцію та розрахунок накопичення рекомбінантного білка GFP при гель-електрофорезі в денатуруючих умовах проводили тільки для варіантів без додавання ацетосирингону. Рівень експресії для усіх експериментів виражали в кількості накопиченого білка у г/кг сирової речовини рослинної біомаси.

Показано, що рівні накопичення білка GFP при інфільтрації рослин бактеріальною суспензією з $OD_{600} = 0,8, 1,0, 1,5$ статистично не відрізнялись та були найвищими зі всіх досліджуваних концентрацій, і становили $0,2935 \pm 0,0082, 0,2970 \pm 0,0092, 0,2885 \pm 0,0095$ г/кг сирової речовини рослинної біомаси відповідно. Інфільтрація рослин суспензією з оптичною густиною $OD_{600} = 0,2, 0,4, 2,0$, призводить до статистично достовірного зменшення накопичення рекомбінантного білка, $0,1540 \pm 0,0072, 0,1529 \pm 0,0067$ та $0,2225 \pm 0,0145$ г/кг сирової речовини рослинної біомаси відповідно. Більш того, статистичної різниці концентрації рекомбінантного білка при використанні суспензії з $OD_{600} = 0,2$ та $OD_{600} = 0,4$ нами не виявлено. Встановлено, що рівень накопичення рекомбінантного білка був в 2,22 рази вищий при використанні бактеріальної суспензії з оптичною густиною рівною 1,0 (OD_{600}) порівняно із бактеріальною суспензією з густиною рівною 0,2.

Досліджено рівень експресії репортерного гена *gfp* використовуючи конструкції з різними промоторними послідовностями, а саме: 35S промотор вірусу мозаїки цвітної капусти (35SCaMV P) (pSPV2301), подвійний 35SCaMV P (D35SCaMV P, pSPV2303), промотори генів *rbcslb* (pSPV2302) та *rbcslb* (pSPV2304), що кодують різні домени малої субодиниці рибулозобісфосфаткарбоксилази (RuBisCo), виділених з різущки Таля *Arabidopsis thaliana* L. (*A. thaliana*), промотори генів *lhb1b1* (pSPV2305) та *lhb1b2* (pSPV2306) з *A. thaliana*, що кодують хлорофіл a-b зв'язуючі білки. Окрім промоторів, генетичні конструкції були доповнені 5'Ω TMV та термінатором 35S T.

Найвищий рівень експресії гена *gfp*, $0,3045 \pm 0,0097$ г/кг сирової речовини рослинної біомаси, спостерігали при використанні промотора D35SCaMV P, що в 5,41 разів вище ніж за промотора LHB1B1 P – $0,0562 \pm 0,0089$ г/кг сирової

речовини рослинної біомаси. При використанні промоторів 35S CaMV P, RbcS1B P, RbcS2B P, LHB1B2 P рівні накопичення рекомбінантного білка становили $0,2135 \pm 0,0079$, $0,0625 \pm 0,0105$, $0,1255 \pm 0,0125$ та $0,1135 \pm 0,0156$ г/кг сирої речовини рослинної біомаси відповідно. Рівні експресії гена *gfp* статистично не відрізнялись між собою при використанні RbcS1B P та LHB1B1 P, а також LHB1B2 P та RbcS2B P. Статистична різниця рівнів накопичення рекомбінантного білка GFP встановлено для промоторів 35S CaMV P, D35SCaMV P, LHB1B2 P/RbcS2B P, RbcS1B P/LHB1B1 P.

Проаналізовано рівень експресії гена *gfp* в залежності від локалізації рекомбінантного білка в різних клітинних та позаклітинних компартментах. Для цього в генетичних конструкціях використовували: хлоропластний синтетичний сигнальний пептид РубісКо (RbcS, pSPV2309), апопластний сигнальний пептид альфа-амілази 3А рису (RAmy3A, pSPV2307), мітохондріальний сигнальний пептид цитохром С-оксидази дріжджів *Saccharomyces cerevisiae* (ScCoxIV, pSPV2308) та конструкцію без сигнального пептида (pSPV2303), яка забезпечувала локалізацію рекомбінантного білка в цитозолі. Конструкції, окрім сигнальних пептидів, були доповнені промотором D35SCaMV P, 5'Ω TMV, термінатором 35S T та кодуючою послідовністю гена *gfp*. При розділенні білків за допомогою гель-електрофорезу визначили, що найнижчий рівень експресії спостерігався при використанні апопластного сигнального пептида, а найвищий – без його використання в конструкції. Кількість рекомбінантного білка GFP в апопласті була нижчою в 4,55 разів ніж у цитозолі, і становила $0,0612 \pm 0,0089$ та $0,2785 \pm 0,0078$ г/кг сирої речовини рослинної біомаси відповідно. При використанні мітохондріального або хлоропластного сигнальних пептидів рівень накопичення рекомбінантного GFP становив $0,1334 \pm 0,0149$ та $0,1467 \pm 0,097$ г/кг сирої речовини рослинної біомаси відповідно, при цьому різниця між концентраціями білка була статистично недостовірною.

Встановлено рівень накопичення рекомбінантного білка GFP при використанні різних 5'-послідовностей, що не транскрибуються (5'UTR's): 5'UTR X-вірусу картоплі *Potato Virus X* (5'UTR Ω PVX, pSPV2310), 5'Ω TMV

(pSPV2303), 5'UTR's вірусу огіркової мозаїки *Cucumber Mosaic Virus* (5'UTR CMV1-pSPV2317, 5'UTR CMV2-pSPV2311), 5' UTR гена 2В, який належить до родини генів, що кодує малу субодиницю РубісКо (5'UTR RbcS2В), виділеного із *A. thaliana* (pSPV2312). Також було використано конструкцію, яка не містила 5'UTR (pSPV2318). Окрім послідовностей 5'UTR, конструкції були доповнені промотором D35SCaMV P, термінатором 35S T і кодуючою послідовністю гена *gfp*.

Найнижчий рівень експресії гена *gfp* спостерігали при використанні конструкції, яка не була доповнена 5'UTR. Так, концентрація рекомбінантного білка в цьому випадку становила $0,1312 \pm 0,0078$ г/кг сирової речовини рослинної біомаси, що в 3,2 разів нижче, ніж за використання 5'UTR RbcS2В ($0,4192 \pm 0,0089$ г/кг сирової речовини рослинної біомаси). Необхідно зазначити, що накопичення рекомбінантного білка в останньому випадку було найвищим серед досліджуваних зразків. Вміст GFP при використанні в генетичних векторах 5'UTR Ω PVX, 5'UTR CMV1, 5'UTR CMV2 складав: $0,1830 \pm 0,0094$, $0,3556 \pm 0,0090$, $0,2412 \pm 0,0094$ та $0,3022 \pm 0,0056$ г/кг сирової речовини рослинної біомаси відповідно. Рівні експресії гена *gfp* статистично відрізнялись для всіх варіантів конструкцій, що передбачали використання різних 5'UTR послідовностей.

Нами було проаналізовано рівень експресії гена *gfp* при використанні в генетичних конструкціях різних термінаторів: *g7* виділений з *A. tumefaciens* (Atug7 T, pSPV2313), гена манопінсинтази з *A. tumefaciens* (mas T, pSPV2314), аденозинтрифосфатази (ATPase T, pSPV2315) томатів *Solanum lycopersicum* L., гістона H4 T картоплі *Solanum tuberosum* L. (pSPV2316), термінатор 35S T (pSPV2312). Окрім термінаторів, регуляція експресії гена *gfp* відбувалась за допомогою 5'UTR RbcS2В та промотора D35SCaMV P.

Визначено рівні експресії гена *gfp* залежно від обраного термінатора у векторі. Найвищий рівень експресії спостерігався при використанні термінатора Atug7 T ($0,5585 \pm 0,0120$ г/кг сирової речовини рослинної біомаси) та перевищував його найменше значення ($0,1930 \pm 0,0042$ г/кг сирової речовини рослинної біомаси) за використання термінатора гена гістону H4 T у 2,89 разів.

За використання термінаторів ATPase T, 35S T, mas T рівень накопичення рекомбінантного білка становив: $0,4670 \pm 0,0098$, $0,4060 \pm 0,0084$ та $0,3355 \pm 0,0077$ г/кг сирої речовини рослинної біомаси відповідно. Показано статистичну різницю накопичення рекомбінантного білка GFP для всіх варіантів конструкцій, що передбачали використання різних термінаторів.

В результаті використання в генетичній конструкції послідовності, що кодує білок-інгібітор посттранскрипційного сайленсингу генів P19 вірусу кущистої карликовості томатів (*Tomato Bushy Stunt Virus*, TBSV), було показано підвищення рівня експресії гена *gfp*. Для цього використовували генетичну конструкцію з двома транскрипційними одиницями (pSPV2324), одна з яких включала ген *gfp* під контролем D35SCaMV P промотора, Atug7 T термінатора та 5'UTR RbcS2B, друга – ген *p19* під контролем промотора 35SCaMV P, 5'Ω TMV та термінатора октопінсинтази *A. tumefaciens* L. (ocs T). Як контроль використовували конструкцію pSPV2313, яка несе тільки одну транскрипційну одиницю з геном *gfp*. При розділенні білків методом гель-електрофорезу визначили, що вміст рекомбінантного білка GFP, з використанням гена *p19* в генетичному векторі, складав $0,8624 \pm 0,0085$ г/кг сирої речовини рослинної біомаси, що у 1,46 разів більше ніж у контролі ($0,5897 \pm 0,0078$ г/кг сирої речовини тканин).

Наукова новизна одержаних результатів полягає у комплексному вивченні гетерологічної експресії генів, на прикладі репортерного гена *gfp*, а саме:

- *уперше* проведено порівняльний аналіз експресії репортерного гена *gfp* при транз'єнтній генетичній трансформації рослин виду *Nicotiana rustica* L. за використання різних регуляторних елементів: промоторів – D35SCaMV P, 35SCaMV P, Rbcs2B P, RbcS1b P, LHB1B2 P, LHB1B1 P; сигнальних пептидів – RbcS, RAmy3A, ScCoxIV; 5'- послідовностей, що не транлюються – 5'UTR Ω PVX, 5'Ω TMV, 5'UTR CMV1, 5'UTR CMV2, 5'UTR RbcS2B; термінаторів – H4 T, 35S T, Atug7 T, mas T, ATPase T. Показано, що рівні експресії гетерологічного гена *gfp* відрізнялись в залежності від використаних регуляторних елементів у генетичному векторі;

- *уперше* підтверджено вплив білка-супресора посттранскрипційного сайленсингу генів Р19 TBSV на експресію гетерологічного гена *gfp* при транз'єнтній генетичній трансформації рослин виду *Nicotiana rustica* L.

Удосконалено систему транз'єнтної трансформації рослин виду *Nicotiana rustica* L., за рахунок визначення оптимальної густини агробактеріальної суспензії.

Практичне значення отриманих результатів.

Результати дисертаційної роботи можуть слугувати важливим теоретичним підґрунтям при створенні генетичних векторів для транз'єнтної і стабільної трансформації рослин, підвищення ефективності експресії генів та оптимізації процесів отримання рекомбінантних білків, а також сприяти розвитку перспективних напрямків дослідження процесів регуляції експресії генів.

Результати теоретичних та практичних досліджень застосовуються:

- в практичній діяльності ТОВ «БЕТА НК» у процесі створення генетичних конструкцій для трансформації рослин.

Ключові слова: експресія генів, рекомбінантний білок, *gfp*, *Nicotiana rustica* L., транз'єнтна трансформація.