
ВКЛАД КОНСЕНСУСНОЙ 5'-НЕТРАНСЛИРУЕМОЙ ОБЛАСТИ В ЭФФЕКТИВНОСТЬ ТРАНСЛЯЦИИ ГЕТЕРОЛОГИЧНЫХ ГЕНОВ В РАСТИТЕЛЬНЫХ КЛЕТКАХ

М.Г. Кимиссе¹, К.В. Кабардаева²,
О.А. Гра³, А.А. Тюрин³

¹ Кафедра ботаники, физиологии растений и агробиотехнологии
Российский университет дружбы народов
ул. Миклухо-Маклая, 8/2, Москва, Россия, 117198

²РГАУ — МСХА имени К.А. Тимирязева
ул. Тимирязевская, 49, Москва, Россия, 127550

³ФГБУН Институт физиологии растений им. К.А. Тимирязева РАН
ул. Ботаническая, 35, Москва, Россия, 127276

Результаты биоинформатического анализа позволили выявить следующую закономерность: средний размер 5'-нетранслируемой области (5'-НТО) для большинства генов *A. thaliana* с высоким уровнем экспрессии варьирует от 80 до 120 п. н. со средним содержанием GC 36,5%. На основании результатов выравнивания определена консенсусная 5'-НТО как новый регуляторный элемент, который потенциально может обеспечить высокоэффективную экспрессию и синтез целевого продукта в растениях. Эта последовательность имеет размер 87 п.н. и GC содержание 35,6%. Продемонстрировано, что консенсусная 5'-НТО обеспечивает увеличение накопления бирепортерного белка более, чем на 25%, тем самым выступает в качестве потенциального позитивного регуляторного элемента в эффективности трансляции.

Ключевые слова: растения, 5'-нетранслируемая область, трансляция, экспрессия генов.

Введение. Получение рекомбинантных белков — одна из ключевых задач биотехнологии. Наиболее актуальным направлением в этой области является получение полипептидов, естественный источник которых крайне ограничен — как правило, это различные белки человека и животных, имеющие фармацевтическое значение. Для получения таких важных полипептидов используется их биосинтез в различных гетерологичных системах: трансгенные животные, культура клеток (в равной степени это относится как к эукариотическим, так и прокариотическим организмам), а также трансгенные растения. Применение растений для наработки целевых белков позволяет частично или полностью решить ряд проблем, возникающих при экспрессии гетерологичных генов в других организмах.

Но наряду с этим существует потребность контролировать экспрессию целевых генов в растительных системах более точно, по возможности имея достаточный уровень рекомбинантного белка, при этом не меняя в значительной степени биохимический профиль клеток-продуцентов и обеспечивая стабильность целевых белков и их доступность для выделения. Использование различных промоторов [1; 11] — достаточно хорошо изученный подход, который, однако, решает только проблему достижения определенного уровня мРНК в клетке (а также кор-

ректировка этого процесса в зависимости от локализации в организме и во времени) [9], но не позволяет сколько-нибудь существенно влиять на стабильности мРНК и эффективность трансляции.

Следует отметить, что 5'-НТО является ключевым регуляторным элементом, определяющим эффективность экспрессии гетерологичного гена на этапе инициации трансляции. Данный участок мРНК предшествует стартовому кодону и, как следует из названия, не транслируется в аминокислотную последовательность белка.

5'-НТО может выступать как в качестве самостоятельного регуляторного элемента, так и быть своего рода контейнером для других регуляторных последовательностей.

Обычно в составе 5'-НТО можно отметить следующие структурные элементы: (1) m7G-сар — структурно модифицированный нуклеотид на 5'-конце мРНК, встречающийся у большинства транскриптов и принимающий участие в так называемом кэп-зависимом механизме трансляции, а также осуществляющий защиту молекулы матричной РНК от преждевременного гидролиза нуклеазами; и (2) консенсусная последовательность Козак — мотив окружающий стартовый кодон и обеспечивающий его успешное распознавание рибосомой [6—8; 12].

Помимо этого другие генетические детерминанты, такие как, например, рибопереключатели, также достаточно часто встречаются в составе 5'-НТО. Эти регуляторные элементы обычно представляют собой участки мРНК, состоящие из двух доменов, один из которых отвечает за связывание с низкомолекулярным лигандом, а второй — за ассоциацию с ферментами транскрипционного комплекса. В дополнении в 5'-НТО могут содержаться upstream-открытые рамки считывания — короткие области, ограниченные стартовым и стоп-кодонами [4; 5]. Как правило, данные последовательности крайне негативно влияют на трансляцию, что сказывается на уровне целевого белкового продукта за счет отвлечения рибосом от основного стартового кодона. Однако, с другой стороны, образующиеся пептиды в ряде случаев могут положительно влиять на экспрессию основного гена.

Исходя из всего вышесказанного, 5'-НТО представляется весьма важным регуляторным элементом, способным оказать разнообразное влияние на обеспечение стабильности мРНК и эффективность ее трансляции. Поэтому, на наш взгляд, при разработке кассет для экспрессии трансгенов в растениях важно учитывать состав области, предшествующей стартовому кодону. Поиску консенсусной 5'-НТО для генов растений, на примере *A. thaliana*, и оценки ее влияния на эффективность трансляции гетерологичного гена в растительных клетках и посвящена работа.

Материалы и методы. Бактериальные штаммы. В работе использованы штамм *E. coli* XL1-Blue, JM110 и BL21 (Stratagene, США) и штамм агробактерий GV3101.

Растительный материал. В исследовании использованы растения табака *Nicotiana benthamiana* возраста 6 недель, которые выращивали в теплице, при температуре 25 °С и 16 часовом дне.

Методы молекулярного клонирования. В исследовании использованы стандартные приемы молекулярного клонирования: полимеразная цепная реакция (ПЦР), рестриктазно-лигазный метод [2]. Эндонуклеазы рестрикции, T4 ДНК-лигаза, Taq- и Pfu-ДНК-полимеразы, фосфатазы использовались согласно протоколам фирм-изготовителей (Promega, США; Fermentas, Литва).

Трансформация бактерий. Для трансформации бактерий *E. coli* (штаммы XL1-Blue и BL21) применяли метод температурного шока [2], трансформанты выращивали на среде с ампицилином. Трансформацию *Agrobacterium tumefaciens* осуществляли при помощи электропорации, трансформанты отбирали на среде с карбеницилином.

Агроинфильтрация растений табака *N. benthamiana*. Культуру клеток *A. tumefaciens* штамм GV3101 выращивали в течение ночи при 28 °С в среде 2YT с антибиотиками рифампицин (50 мкг/мл), гентамицин (25 мкг/мл) и антибиотик на агробактериальный вектор карбеницилин (50 мкг/мл). Клетки осаждали при 5000 g в течение 5 минут и ресуспендировали в буфере для агроинфильтрации (10 mM MgSO₄, 10 mM MES pH 5,6). Концентрацию клеток измеряли по величине оптической плотности OD 600, доводили до OD 600 = 0,2. Далее клетки инфильтрировали в нижнюю часть листа шестинедельных растений *N. benthamiana* с помощью шприца без иглы. Агроинфильтрацию индивидуальными штаммами агробактерий проводили не менее, чем в 3-х независимых экспериментах в 5 биологических повторностях. В качестве контроля при проведении экспериментов были использованы растения табака *N. benthamiana*, агроинфильтрированных штаммом, не несущим экспрессионные вектора.

Получение белковых лизатов. Экстракцию суммарных растворимых белков из растительного материала проводили путем растирания свежей или замороженной жидким азотом растительной ткани в 50 mM Tris-HCl (pH = 8,0) буфере (для экстракции 1 г ткани использовали 2—4 мл буфера). Экстракт центрифугировали при 10000 g в течение 10 мин и отбирали надосадочную жидкость.

Анализ белковых лизатов. Количество белка в препаратах определяли по методу Bradford [3], используя Bio-Rad dye reagent (BioRad, США) и БСА (Sigma, США) для построения калибровочной кривой. Активность лихеназы определяли, используя лихенан (Megazyme, Ирландия) в качестве субстрата, время инкубации 10 мин. Определение восстанавливающих сахаров, освобождающихся из субстрата, проводили по методу Вуда и Бхат [13].

Результаты и обсуждение. Гипотеза, положенная в основу представленной работы, заключалась в том, что 5'-НТО генов со сходным уровнем экспрессии с некоторой долей вероятности будут иметь схожие мотивы. Первоначально мы хотели выяснить, как распределяются гены *Arabidopsis thaliana* в зависимости от длины 5'-НТО. Результаты проведенного анализа визуализированы в виде графика, выведенного программным обеспечением FlowGene, представлены на рис. 1.

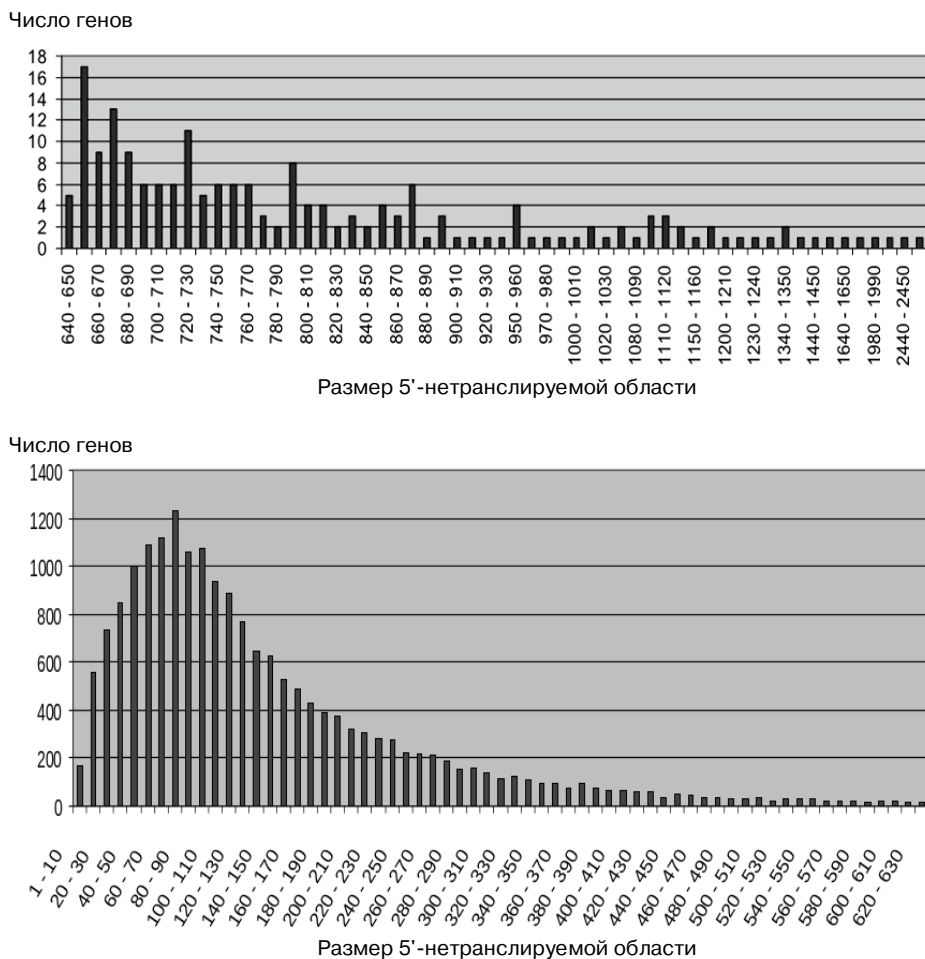


Рис. 1. Распределение генов *A. thaliana* в зависимости от размера 5'-нетранслируемой области

Проведенный анализ показал, что размер 5'-НТО варьирует в широких пределах: минимальный — 1 п.н. и максимальный — 3214 п.н. Как видно из представленных на данном рисунке результатов, большинство генов *A. thaliana* имеет 5'-НТО размером от 50 до 110 п.н. При этом менее 2% генов имеет размер этой области более 500 п.н. Для того, чтобы выяснить есть ли зависимость между размером 5'-НТО и уровнем экспрессии генов (на уровне транскрипции), был проведен анализ с использованием базы данных FlowGene и сопровождающего программного обеспечения. Для анализа этой зависимости были сформированы пять групп генов *A. thaliana* с разным уровнем относительной экспрессии. Согласно такому разделению 42,1%, 46,8%, 10,8%, 0,3% и 0,05% генов арабидопсиса имеют уровень относительной экспрессии от 1 до 100, от 101 до 1000, от 1001 до 10 000, от 10 001 до 20 000 и от 20 001 и выше соответственно.

С использованием программного обеспечения базы данных FlowGene для генов с различным уровнем экспрессии рассчитано распределение их числа в зависимости от размера 5'-нетранслируемой области и их уровнем экспрессии. На рис. 2 представлены полученные в результате анализа данные.

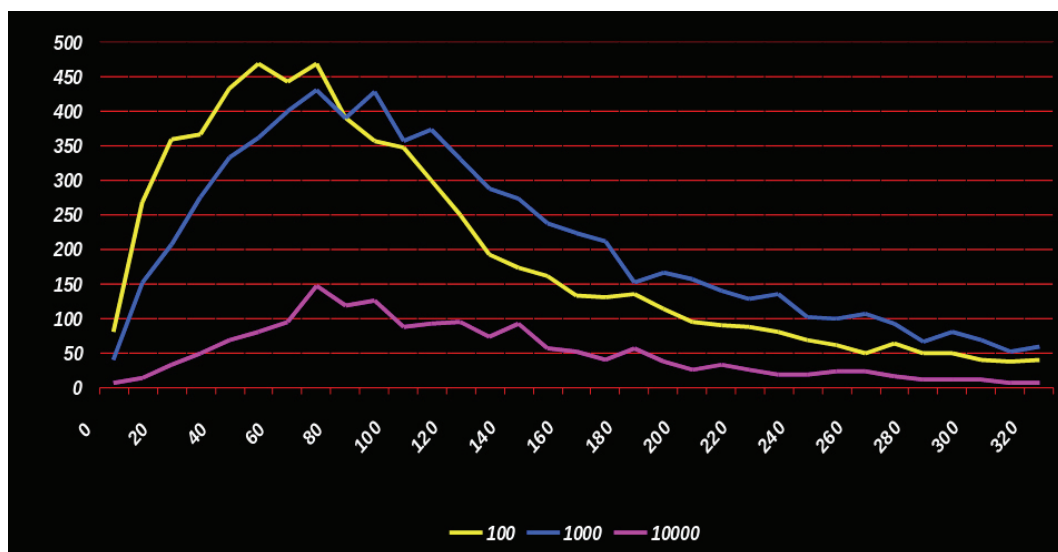


Рис. 2. Распределение генов *A. thaliana* в зависимости от размера 5'-нетранслируемой области в пределах 1—310 п.н. и уровня экспрессии от 1 до 100, от 101 до 1000, от 1001 до 10 000 относительных единиц

Проведенный анализ позволил установить следующие закономерности:

— размер 5'-НТО для генов с уровнем относительной экспрессии от 1 до 100 варьирует в широких пределах от 1 п.н. до 1600 п.н., при этом для большинства генов этой группы характерен размер 5'-НТО от 20 до 180 п.н.;

— размер 5'-НТО для генов с уровнем относительной экспрессии от 101 до 1000 варьирует в широких пределах от 1 п.н. до 1400 п.н., при этом для большинства генов этой группы размер характерен 5'-НТО от 50 до 130 п.н.;

— размер 5'-НТО для генов с уровнем относительной экспрессии от 1001 до 10 000 варьирует в менее широких пределах от 1 п.н. до 1000 п.н., при этом для большинства генов этой группы характерен размер 5'-НТО от 60 до 140 п.н.;

— размер 5'-НТО для генов с уровнем относительной экспрессии от 10 001 до 20 000 варьирует в пределах от 30 п.н. до 210 п.н., при этом размер 5'-НТО от 80 до 160 п.н. характерен для большинства генов этой группы.

Далее был рассчитан средний размер 5'-нетранслируемой области для каждой группы генов *A. thaliana*, который составил: 128,7 п.н., 156,3 п.н., 139,2 п.н., 127 п.н. и 97 п.н. для группы с относительным уровнем экспрессии от 1 до 100, от 101 до 1000, от 1001 до 10 000, от 10 001 до 20 000 и от 20 001 до 40 000 соответственно. Т.е. отмечено уменьшение размера этой области с увеличением уровня экспрессии генов.

Как уже отмечено ранее, предполагается, что важную роль в уровне экспрессии может играть не только размер 5'-НТО генов, а и ее нуклеотидный состав, а именно содержание GC и AT нуклеотидов. Для того, чтобы выяснить, есть ли зависимость между GC содержанием в 5'-нетранслируемой области и уровнем экспрессии генов (на уровне транскрипции), у *A. thaliana* был проведен анализ с использованием базы данных FlowGene и сопровождающего программного обеспечения.

Для анализа этой зависимости были использованы, сформированные ранее три основные группы генов (в которые входит большинство генов) с разным уровнем относительной экспрессии (максимальное значение уровня относительной экспрессии): от 1 до 100, от 101 до 1000, от 1001 до 10 000.

Проведенный анализ позволил установить следующие закономерности: для всех трех групп генов по уровню экспрессии наблюдается картина, соответствующая нормальному распределению (в общих чертах). В группе, включающей гены с уровнем экспрессии от 1 до 100 относительных единиц, распределение с максимумом, составляющим 34—35%, также имеет незначительно отклоняющиеся пики для генов с процентным содержанием гуанина и цитозина 33 и 37%. Во второй группе (уровень экспрессии — 101—1000 отн. ед.) распределение числа генов в зависимости от содержания GC более гладкое, пик приходится на 37%. Для третьей группы максимум соответствует точке 36%, также имеются пики с содержанием GC 37 и 38%. Во всех трех группах практически отсутствуют гены с содержанием гуанина и цитозина ниже 20% и выше 55%.

Из вышесказанного можно заключить, что для большинства генов GC содержание в 5'-нетранслируемой области в среднем составляет около 36,5%.

Таким образом, результаты биоинформатического анализа позволили выявить следующую закономерность: средний размер 5'-нетранслируемой области для большинства генов *A. thaliana* с относительно высоким уровнем экспрессии варьирует от 90 до 200 п.н. со средним содержанием GC 36,5%. Следует отметить, что зачастую исследователи клонируют и апробируют 5'-НТО конкретного гена организма, что позволяет определить ее функциональную значимость на эффективность трансляции, однако не дает ответ на вопрос о том, что именно наличие 5'-НТО в последовательности гена влияет на эффективность трансляции растительных генов.

Для того, чтобы выяснить есть ли взаимосвязь между уровнем экспрессии гена (на этапе трансляции) и наличием или отсутствием у него 5'-НТО, мы провели теоретические и экспериментальные исследования, результаты которых изложены ниже. Для этого, первоначально необходимо было найти консенсусную последовательность 5'-НТО, характерную для генов растений с высоким уровнем транскрипции. С этой целью были использован следующий алгоритм для теоретического предсказания консенсусной 5'-НТО:

1) выбраны по 1000 случайных последовательностей в группах генов с разным уровнем транскрипции до 100, до 1000, до 10 000 и имеющим размер от 100 до 199 п.н.;

2) в каждой группе последовательностей за счет множественного выравнивания определены мотивы, проявляющие консервативность и имеющие среднее содержание GC до 37%;

3) полученные мотивы объединены и определен размер последовательностей (87 п.н.) для последующего анализа и выявления консенсусной последовательности;

4) использовали программу web-LOGO для визуализации консенсусной последовательности.

В результате проведенных теоретических работ была определена 5'-НТО как новый регуляторный элемент, который потенциально может обеспечить высоко-

эффективную трансляцию, и, как следствие, высокий уровень целевого белкового продукта в растениях (рис. 3). Эта последовательность имеет размер 87 п.н. и GC содержание 35,6%.

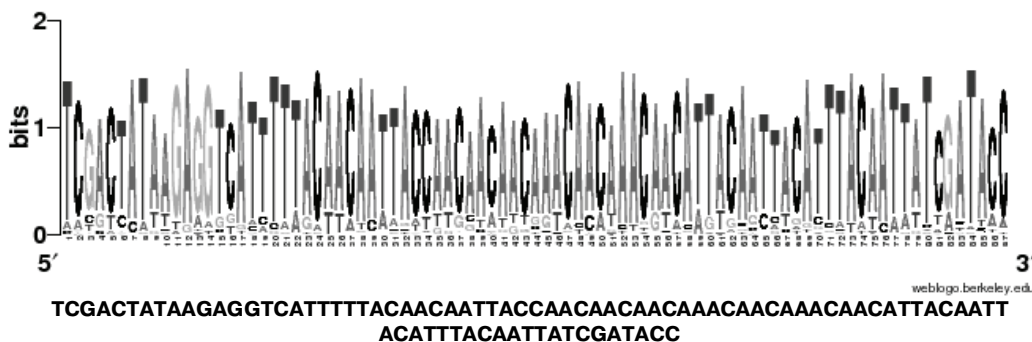


Рис. 3. Консенсусная последовательность 5'-нетранслируемой области выбранная для дальнейших исследований

Для того, чтобы выяснить как консенсусная 5'-НТО, выведенная на основании предсказаний, сказывается на эффективности трансляции, была проведена серия экспериментов.

В качестве векторов для экспрессии были использованы модульные вектора серии *pPGG* и *pVIG-T*, сконструированные в группе функциональной геномики ранее [1]. Выбор этих векторов для нашего исследования обусловлен следующими их преимуществами: вектор *pPGG* содержит оптимальное окружение иницирующего кодона целевого гена для его эффективной трансляции (в положении -1 — А, а в положении +4 — G).

Введение именно этих нуклеотидов основано на том, что окружение иницирующего AUG кодона является важным фактором для эффективной трансляции целевого гена в растениях. Многочисленными исследованиями подтверждена важность нуклеотидов в положениях -1 и +4 в этой последовательности в эффективности трансляции генов. Известно, что оптимальный контекст для инициации трансляции в клетках животных и растений, по-видимому, один и тот же и представлен следующей последовательностью: *AACAATGGC*. Иницирующие кодоны в контексте, отличающемся от оптимального, узнаются рибосомами менее эффективно и допускают их прохождение до следующего иницирующего кодона.

Вектор *pVIG-T*, сконструированный на основе плазмиды *pICH6692*, имеет размер 6342 п.н. В Т-ДНК область вектора включены следующие последовательности: селективный ген *NPTII*, экспрессия которого контролируется *pNos* промотором и *Nos* терминатором; последовательность полилинкера, включающая уникальные сайты с интегрированным между ними случайной последовательности ДНК, для простой процедуры переноса экспрессионного модуля из вектора *pPGG*. Этот вектор сконструирован для обеспечения стабильной экспрессии гетерологичных генов в растениях.

Поскольку селективный ген не является необходимым для транзIENTНОЙ экспрессии, в связи с тем, что не нужно обеспечивать отбор первичных трансформантов на селективной среде, сконструирован вектор *pVIG-S* на основе плазмиды

pICN6692, который имеет размер 4570 п.н. В Т-ДНК область вектора включены следующие последовательности: ген p19, экспрессия которого контролируется промотором ТСТР и *osc*-терминатором; последовательность полилинкера, включающая уникальные сайты с интегрированным между ними случайной последовательности ДНК, для простой процедуры переноса экспрессионного модуля из вектора pPGG.

Следует отметить, что промотор ТСТР, используемый для обеспечения экспрессии гена p19, как было показано, обеспечивает высокий уровень конститутивной экспрессии целевого гена [10]. Этот вектор сконструирован для обеспечения транзientной экспрессии гетерологичных генов в растениях и дает возможность использовать только один штамм бактерий для агробактериальной инфильтрации, в отличие от других векторов, для которых необходимо использовать процедуру котрансформации двумя штаммами агробактерий, один из которых несет вектор с геном p19, а другой — вектор с целевым геном. Первоначально методом ПЦР синтезирован фрагмент ДНК, соответствующий консенсусной последовательности 5'-НТО. Далее этот фрагмент был интегрирован в базовый вектор серии pPGG в область между промотором и стартовым кодоном бирепортерного гена *licB-gfp* (далее — LG), в котором ген термостабильной лихеназы (*licB*) и зеленого флуоресцентного белка (*gfp*) имеют транскрипционно-трансляционное слияние. Затем, используя процедуры молекулярного клонирования, был получен экспрессионный вектор для транзientной экспрессии в растениях pVIG-T-87-UTR-LG. На рис. 4 представлены схемы использованных векторов, которые несут исследуемый регуляторный элемент.

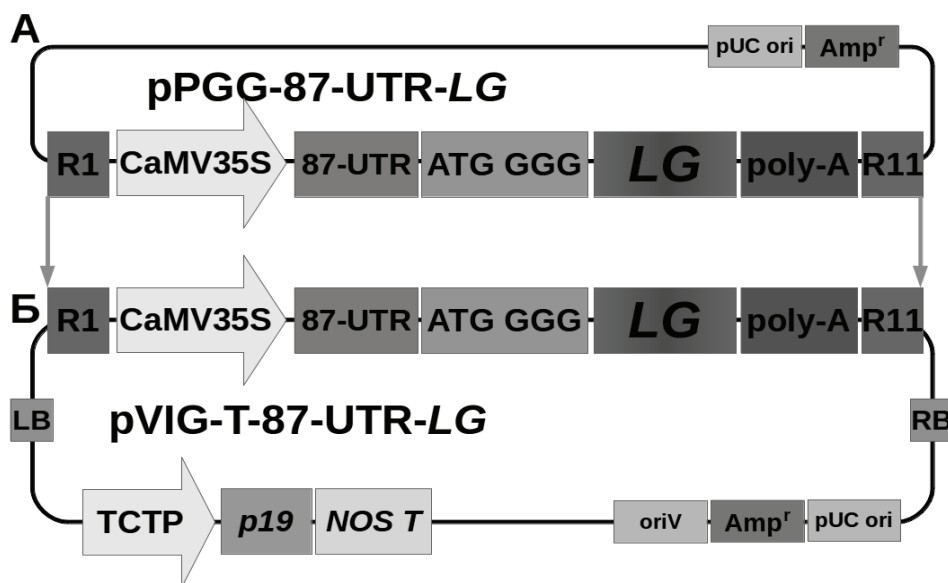


Рис. 4. Схема векторов:

А — pPGG-87-URT-LG; В — pVIG-T-87-URT-LG. CaMV35S — 35 промотор вируса мозаики цветной капусты. ТСТР — промотор гена трансляционно контролируемого опухолевого белка (translationally controlled tumor protein). p19 — ген p19 из вируса *Symbidium ringspot* (CymRSV). LB и RB — левый и правый бордеры соответственно. R1–R11 — сайты рестрикции для клонирования целевого гена в состав экспрессионной кассеты и для ее переноса в окончательный растительный вектор. pUC ori oriV — точка начала репликации для *E. coli* и *A. tumefaciens*. Amp^r — ген устойчивости к ампицилину. polyA — сигнал полиаденилирования. 87-UTR — синтетическая консенсусная 5'-НТО

Полученным вектором трансформированы агробактерии (*Agrobacterium tumefaciens*) штамма GV3101, которые затем использованы для агроинфильтрации 6-недельных растений табака *Nicotiana benthamiana*. Первоначально эффективность процесса агроинфильтрации оценивали по флуоресценции зеленого флуоресцентного белка, который экспрессируется в составе бирепортерного белка. Через 4 суток после агроинфильтрации из зараженных частей листьев готовили белковые экстракты. Для сравнительного анализа эффективности регуляторного элемента был определен уровень накопления бирепортерного белка в листьях растений табака *N. benthamiana* при агроинфильтрации. В качестве контроля использовали экспрессионный вектор pVIG-T-LG, полученный в группе функциональной геномики ранее. Оценка уровня накопления рассчитывали по активности лихеназы, как части бирепортерного белка, в перерасчете на суммарный растворимый белок. На рис. 5 представлены суммируемые результаты тестирования эффективности консенсусной 5'-НТО как регуляторного элемента трансляции.

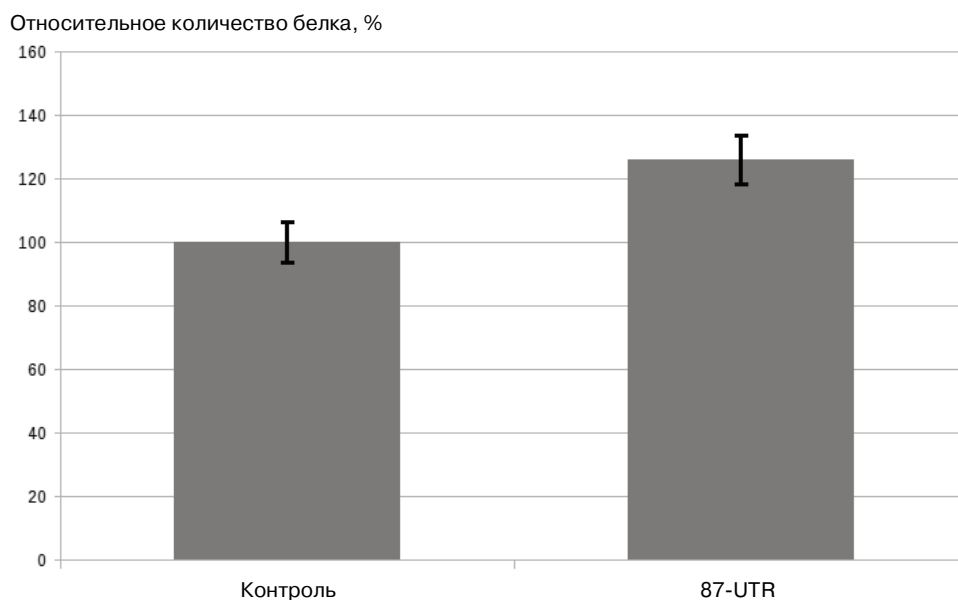


Рис. 5. Уровни накопления репортерного белка при использовании синтетической консенсусной 5'-НТО. Относительно контроля (вектор с той же репортерной системой, но без 87-UTR — pVIG-T-LG). Контроль — pVIG-T-LG, 87-UTR — pVIG-T-87-UTR-LG

Как видно из полученных результатов, наличие в составе экспрессионного вектора pVIG-T-87-UTR-LG синтетической консенсусной последовательности для 5'-НТО бирепортерного гена обеспечило увеличение накопления бирепортерного белка более чем на 25% по сравнению с контролем (в данном случае экспрессионный вектор pVIG-T-LG, в котором бирепортерный ген не содержит 5'-НТО). Полученные результаты позволили предположить, что синтетическая консенсусная последовательность, обозначенная нами как 87-UTR, выступает в качестве потенциального регуляторного элемента эффективности трансляции.

Продемонстрированный рост уровня экспрессии целевого гена под дополнительным управлением синтетической консенсусной 5'-НТО является, тем не менее, лишь эмпирическим фактом, не имеющим пока молекулярно-генетического объяснения. Наличие 5'-нетранслируемой области обычно не предусматривается в составе экспрессионных векторов. Трансляция подавляющего большинства эукариотических генов происходит по так называемому сканирующему (кэп-зависимому) механизму, суть которого состоит в сканировании малой субъединицей рибосомы последовательности мРНК в направлении 5'–3' от 5'-кэпа к стартовому кодону — AUG. Матричная РНК небольшого числа (около 70 генов эукариот) генов транслируется по механизму внутренней инициации (кэп-независимому, IRES-зависимому), в этом случае рибосома обнаруживает участок мРНК, называемый сайтом внутренней посадки рибосомы — IRES. Реже встречается еще один механизм — «бесфазное блуждание».

Таким образом, из всего этого мы можем заключить, что для эффективной экспрессии практически всех (с незначительными исключениями) требуется как минимум 5'-кэп.

Тем не менее, как упоминалось ранее в рамках данного раздела, как правило, лишь крайне незначительное количество генов не имеет 5'-НТО, что, как мы полагаем, связано с определенными функциями этой генетической детерминанты в процессе реализации генетической информации. Мы также предполагаем, что помимо (что было показано, например, на человеке) повышения стабильности матричной РНК, определения ее компартиментализации и некоторых других функций 5'-НТО повышает вероятность связывания рибосомы и трансляционных факторов с молекулой мРНК (в рамках сканирующего механизма, поскольку для механизма внутренней инициации наличие 5'-нетранслируемой области является критичным, т.к. в данном случае она выступает в качестве контейнера для IRES), выступая своего рода платформой для их посадки. С другой стороны, эффективная экспрессия генов при отсутствии данной регуляторной области, на наш взгляд, может быть объяснена большим количеством транскрипта. Однако если рассматривать это явление в рамках предложенного комплексного подхода к точной регуляции экспрессии, это не всегда может являться оптимальным решением задачи наработки целевого белка.

Таким образом, можно заключить, что использование в качестве дополнительного регуляторного элемента синтетической консенсусной 5'-НТО приводит к увеличению уровня накопления рекомбинантного белка на 25%. В значительной степени это связано с наличием ряда регуляторных элементов в составе 5'-НТО мРНК, которые могут определять как характер связывания с различными регуляторными белками, так и уровень стабильности транскрипта.

Также наличие 5'-НТО является немаловажным для посадки рибосомы на молекулу мРНК. Помимо всего прочего следует отметить возможность предсказания функциональной активности регуляторных мотивов на основании определения частоты их встречаемости в геноме. Но, на наш взгляд, результаты, полученные

с применением программного обеспечения базы данных FlowGene, можно улучшить, интегрируя в синтетические 5'-нетранслируемые области не только часто встречаемые мотивы, но и конкретные регуляторные элементы из числа тех, что встречаются в нативных в 5'-НТО, что может послужить основой для дальнейших исследований в этой области практической биотехнологии растений. Предложенный подход, на наш взгляд, может быть использован для получения консенсусных последовательностей также и для других видов растений, как двудольных, так и однодольных.

ЛИТЕРАТУРА

- [1] Вячеславова А.О. Серия модульных векторов для стабильной и транзиторной экспрессии гетерологичных генов в растениях / А.О. Вячеславова, О.Н. Мустафаев, А.А. Тюрин, Х.Р. Шимшилашвили, И.Н. Бердичевец, Д.М. Шаяхметова, М.А. Голденков, В.С. Фадеев, Ю.В. Шелудько, И.В. Голденкова-Павлова // *Генетика*. 2012. Т. 48. № 9. С. 1046—1056.
- [2] *Maniatis T., Fritsch E., Sambrook J.* Молекулярное клонирование. М.: Мир, 1984.
- [3] *Bradford M.A.* Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding // *Analytical Biochemistry*. 1976. Vol. 72. P. 248—254.
- [4] *Calvo S.E., Pagliarini D.J., Mootha V.K.* Upstream open reading frames cause wide spread reduction of protein expression and are polymorphic among humans // *Proceedings of the National Academy of Sciences of the United States of America*. 2009. Vol. 106. P. 7507—7512.
- [5] *Iacono M., Mignone F., Pesole G.* uAUG and uORFs in human and rodent 5'untranslated mRNAs // *Gene*. 2005. Vol. 349. P. 97—105.
- [6] *Kozak M.* Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes // *Proceedings of the National Academy of Sciences of the United States of America*. 1990. Vol. 87. P. 8301—8305.
- [7] *Kozak M.* Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo // *Nature*. 1984. Vol. 308. P. 241—246.
- [8] *Kozak M.* Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes // *Cell*. 1986. Vol. 44. P. 283—292.
- [9] *Mann D.G.J., King Z.R., Liu W., Joyce B.L., Percifield R.J., Hawkins J.S., LaFayette P.R., Artelt B.J., Burris J.N., Mazarei M., Bennetzen J.L., Parrott W.A., Neal Stewart C.J.* Switchgrass (*Panicum virgatum* L.) polyubiquitin gene (PvUbi1 and PvUbi2) promoters for use in plant transformation // *BMC Biotechnol.* 2011. Vol. 11. № 11. P. 74.
- [10] *Masura S.S., Ahmad P.G.K., Ti L.L.E.* Isolation and characterization of an oil palm constitutive promoter derived from a translationally control tumor protein (TCTP) gene // *Plant Physiology and Biochemistry*. 2011. Vol. 49. P. 701—708.
- [11] *Park S.H., Yi N., Kim Y.S., Jeong M.H., Bang S. W., Choi Y.D., Kim J.K.* Analysis of five novel putative constitutive gene promoters in transgenic rice plants // *Journal of Experimental Botany*. 2010. Vol. 61. № 9. P. 2459—2467.
- [12] *Pelletier J., Sonenberg N.J.* Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency // *Cell*. 1985. Vol. 40. P. 515—526.
- [13] *Wood T.M., Bhat K.M.* Methods for measuring cellulase activities // *Methods in Enzymology*. 1988. Vol. 160. P. 87—112.

CONTRIBUTION OF CONSENSUS 5'-UNTRANSLATED REGION TO THE TRANSLATIONAL EFFICIENCY OF HETEROLOGOUS GENES IN PLANT CELLS

M.G. Quimisse¹, K.V. Kabardaeva²,
O.A. Gra³, A.A. Tyurin³

¹Department of botany, plant physiology and agrobiotechnology
Peoples' Friendship University of Russia
Miklukho-Maklaya str., 8/2, Moscow, Russia, 117198

²RSAU — MTAА named after K.A. Timiryazev
Timiryazevskaya str., 49, Moscow, Russia, 127550

³FSBIS K.A. Timiryazev institute of plant physiology RAS
Botanicheskaya str., 35, Moscow, Russia, 127276

The results of bioinformatics analysis allowed to reveal the following consistent pattern: the average length of the 5'-untranslated region (5'-UTRs) for most *A. thaliana*'s genes with high expression levels range from 80 to 120 bp, with an average GC content of 36,5%. Based on alignment results was defined motive in 5'-UTRs, as a new regulatory element, which could potentially provide highly efficient expression and synthesis of the target product in plants. This sequence has a length of 87 bp and GC content of 35.6%. It was demonstrated that consensus sequence of 5'-UTRs increased accumulation of the bi-reporter protein more than 25%, by that acting as a potential positive regulatory element on translational efficiency.

Key words: plants, 5'-untranslated region, translation, gene expression.

REFERENCES

- [1] Vyacheslavova A.O. Seriya modul'nykh vektorov dlya stabil'noy i tranzientnoy ekspressii geterologichnykh genov v rasteniyakh [A Series of modular vectors for providing stable and transient expression of heterologous genes in plants] / A.O. Vyacheslavova, O. Mustafaev, A.A. Tyurin, H.R. Shimshilashvili, I.N. Berdichevets, D.M. Shayakhmetova, M.A. Goldenkov, V.S. Fadeyev, YU.V. Shelud'ko, I.V. Goldenkova-Pavlova. *Genetika* [Genetics], 2012. Vol. 48. N 6. P. 1046—1056.
- [2] Maniatis T., Fritsch E., Sambrook J. *Molekuljarnoe klonirovanie* [Molecular Cloning]. Moscow, Mir Publ., 1984.
- [3] Bradford M.A. Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 1976. Vol. 72. P. 248—254.
- [4] Calvo S.E., Pagliarini D.J., Mootha V.K. Upstream open reading frames cause wide spread reduction of protein expression and are polymorphic among humans. *Proceedings of the National Academy of Sciences of the United States of America*, 2009. Vol. 106, P. 7507—7512.
- [5] Iacono M., Mignone F., Pesole G. uAUG and uORFs in human and rodent 5'untranslated mRNAs. *Gene*, 2005. Vol. 349. P. 97—105.
- [6] Kozak M. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proceedings of the National Academy of Sciences of the United States of America*, 1990. Vol. 87. P. 8301—8305.
- [7] Kozak M. Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo. *Nature*, 1984. Vol. 308. P. 241—246.
- [8] Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, 1986. Vol. 44. P. 283—292.

- [9] Mann D.G.J., King Z.R., Liu W., Joyce B.L., Percifield R.J., Hawkins J.S., LaFayette P.R., Artelt B.J., Burris J.N., Mazarei M., Bennetzen J.L., Parrott W.A., Neal Stewart C.J. Switchgrass (*Panicum virgatum* L.) polyubiquitin gene (PvUbi1 and PvUbi2) promoters for use in plant transformation. *BMC Biotechnol*, 2011. Vol. 11. N 11. P. 74.
- [10] Masura S.S., Ahmad P.G. K., Ti L.L.E. Isolation and characterization of an oil palm constitutive promoter derived from a translationally control tumor protein (TCTP) gene. *Plant Physiology and Biochemistry*, 2011. Vol. 49. P. 701—708.
- [11] Park S.H., Yi N., Kim Y.S., Jeong M.H., Bang S.W., Choi Y.D., Kim J.K. Analysis of five novel putative constitutive gene promoters in transgenic rice plants. *Journal of Experimental Botany*, 2010. Vol. 61. N 9. P. 2459—2467.
- [12] Pelletier J., Sonenberg N.J. Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell*, 1985. Vol. 40. P. 515—526.
- [13] Wood T.M., Bhat K.M. Methods for measuring cellulase activities. *Methods in Enzymology*, 1988. Vol. 160. P. 87—112.