



## КОЛИЧЕСТВЕННОЕ СРАВНЕНИЕ ЭКСПРЕССИИ ГЕНОВ ПРИ ИХ СОЕДИНЕНИИ С ПОМОЩЬЮ IRES ИЛИ КОДИРУЮЩЕЙ ПОСЛЕДОВАТЕЛЬНОСТИ 2А-ПЕПТИДА ТЕШОВИРУСА СВИНЕЙ В БИЦИСТРОННОМ ВЕКТОРЕ

© 2013 г. А. И. Кузьмич<sup>#</sup>, А. В. Введенский, Е. П. Копанцев, Т. В. Виноградова

ФГБУН Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова  
Российской академии наук, Москва, 117997, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 15.01.2013 г. Принята к печати 29.01.2013 г.

В биотехнологии часто требуется осуществить одновременную экспрессию нескольких целевых генов. С этой целью активно разрабатываются мультицистронные векторы, кодирующие несколько белков. Наиболее часто в коммерчески доступных векторах такого типа используют различные варианты внутреннего сайта посадки рибосом вируса энцефаломиокардита (*IRES EMCV*). Однако многие исследователи считают более перспективным использование бицистронных векторов на основе последовательностей, кодирующих саморасщепляющиеся 2А-пептиды. В работе проведено сравнение эффективности экспрессии генов в клетках, трансфицированных бицистронными векторами на основе *IRES EMCV* или нуклеотидной *P2A*-последовательности, соответствующей 2А-пептиду тешовируса-1 свиней. Эффективность экспрессии определена в трех клеточных линиях млекопитающих посредством измерения уровня коэкспрессии генов, кодирующих флуоресцентные белки RFP и EGFP, связанных друг с другом через *IRES* или *P2A*-последовательность. Более высокий уровень экспрессии трансгенов обнаруживался в клетках, трансфицированных генетической конструкцией на основе *P2A*-последовательности.

**Ключевые слова:** бицистронный вектор, *IRES*, *EMCV*, 2А-пептид, тешовирус свиней.

DOI: 10.7868/S0132342313040131

### ВВЕДЕНИЕ

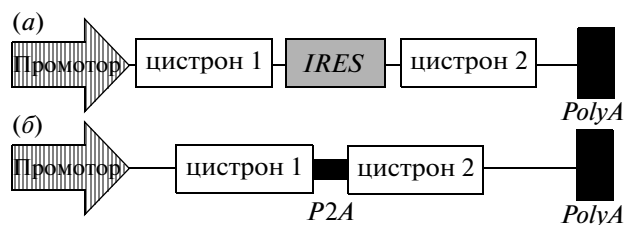
Би- и мультицистронные векторы часто используются в различных биологических исследованиях, в том числе для практических целей [1]. Такие векторы обеспечивают коэкспрессию генов в относительно постоянных стехиометрических соотношениях [1]. Для создания мультицистронных векторов существует множество способов, каждый из которых имеет свои преимущества и недостатки [1]. В настоящее время наиболее популярны бицистронные векторы с сайтом внутренней посадки рибосомы вируса энцефаломиокардита (*IRES EMCV*), схематически представленные на рис. 1. Это связано с высокой эффективностью

внутренней инициации трансляции, обеспечиваемой *IRES EMCV*, а также с тем, что данный элемент открыт давно и достаточно хорошо изучен [2].

В лабораторной практике используются различные варианты *IRES EMCV*. Исходно было показано, что для внутренней инициации трансляции достаточен *IRES*, соответствующий участку 230–834 п.о. геномной РНК EMCV [3, 4]. В генетических векторах используются варианты *IRES* длиной от 430 до 600 п.о. [5]. Модификации нуклеотидной последовательности *IRES EMCV* могут

Сокращения: НТО – нетранслируемая область; РС – рамка считывания; ЭПР – эндоплазматический ретикулум; CMV – промотор цитомегаловируса; EGFP – зеленый флуоресцентный белок; *IRES* – внутренний сайт посадки рибосомы; NMD – распад мРНК, вызванный присутствием преждевременного терминационного кодона (nonsense-mediated decay); *P2A* и *P2A* – пептид 2А тешовируса-1 свиней и кодирующая его последовательность; RFP – красный флуоресцентный белок.

<sup>#</sup> Автор для связи (тел.: +7 (495) 330-69-92; факс +7 (495) 330-46-54-38; эл. почта: akrubik@gmail.com).



**Рис. 1.** Схема строения бицистронных векторов на основе *IRES* (а) и *P2A*-последовательности (б). *PolyA* – сигнал полиаденилирования.

приводить к существенному подавлению внутренней инициации трансляции [5, 6]. Как правило, при создании бицистронных векторов модификациям подвергаются концевые участки *IRES*. При этом укорочение последовательности *IRES* с 5'-конца на 100–200 п.о. практически не сказывается на его активности [7]. Однако нередко изменения нуклеотидной последовательности затрагивают 3'-концевой участок *IRES*. Уровень инициации трансляции резко падает при удалении инициаторного кодона, с которого начинается синтез белка-предшественника у вируса: простое удаление инициаторного кодона или его замена на нуклеотидную последовательность сайта рестрикции *HindIII* снижало экспрессию *IRES*-зависимого цистрона в десятки раз [5, 6].

Эффективность работы *IRES* увеличивается при внедрении вслед за инициаторным кодоном трех кодонов, взятых из генома вируса *EMCV*. Данная модификация позволяет уменьшить влияние белок-кодирующей последовательности на внутреннюю инициацию трансляции, но при этом образующийся белок будет нести четыре дополнительных аминокислотных остатка на *N*-конце [8].

Векторы на основе *IRES* обладают рядом недостатков. В большинстве случаев внутренняя инициация трансляции при помощи *IRES* происходит менее эффективно, чем кэп-зависимая инициация, что приводит к разному уровню синтеза белков, кодируемых первой и второй рамками считывания (PC) [9]. Бицистронная мРНК подвержена деградации в большей степени в силу того, что терминация трансляции первой PC происходит на значительном удалении от 3'-нетранслируемой области (НТО). Терминирующий кодон в середине молекулы может запускать деградацию мРНК по механизму NMD (Nonsense-mediated decay) [10].

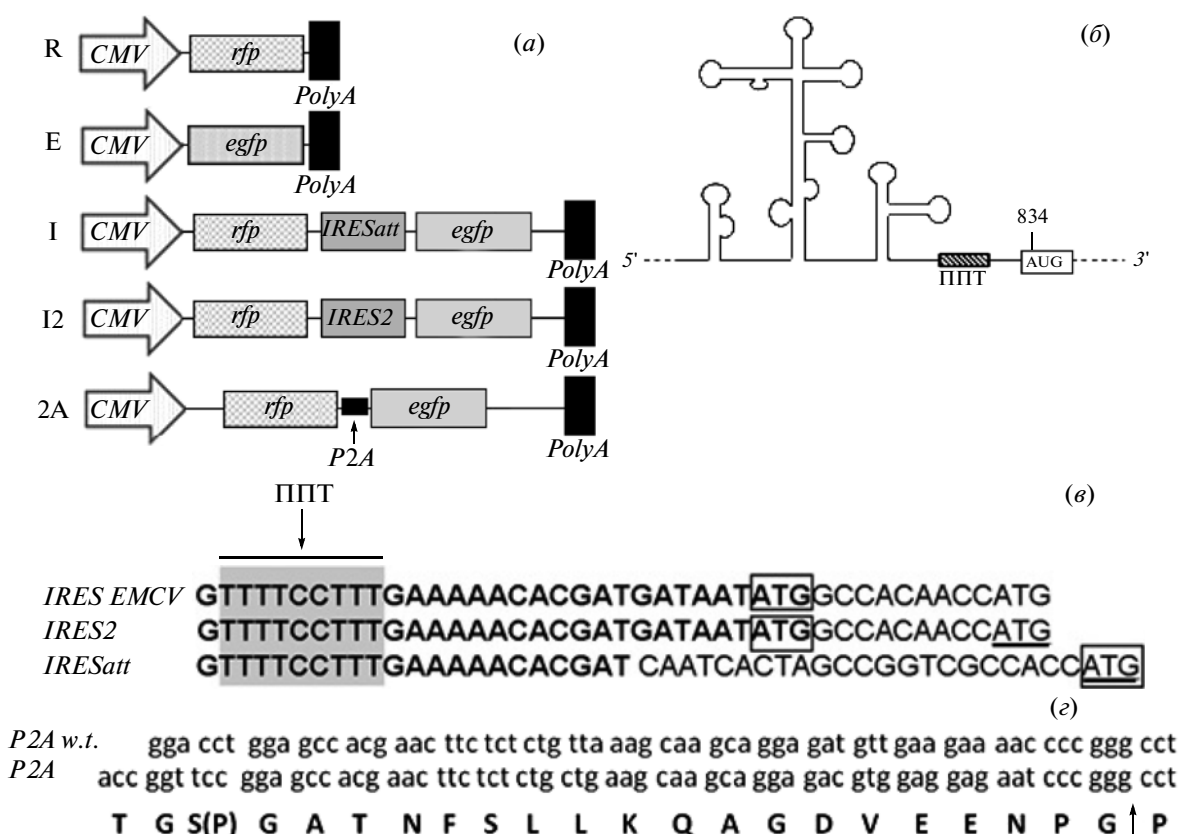
Другой перспективный подход создания мультицистронных векторов основан на использовании в гибридных белках 2A-пептидов или CHYSEL (*cis*-acting hydrolase element) элементов [11]. Эти пептиды кодируются 2A-последовательностями, разделяющими структурные и неструктурные гены в геноме многих пикорнавирусов [12], и имеют размер около 20 а.о. Их *N*-концевые аминокислотные участки сильно различаются у разных представителей пикорнавирусов, однако на *C*-конце у всех 2A-пептидов содержится консервативный фрагмент Asn-Pro-Gly-Pro, необходимый для процессинга 2A-пептида [13]. Этот процесс происходит котрансляционно и состоит в том, что пептидная связь между *C*-концевыми остатками Gly и Pro 2A-пептида не образуется [13]. В результате этого с одной PC образуется два разных белковых продукта с примерным соотношением 1 : 1 [14], при этом первый белок несет около 19 дополнительных аминокислотных остатков на *C*-конце, а второй — остаток Pro на *N*-конце.

Высокий уровень коэкспрессии трансгенов, возможность коэкспрессии трех и более цистронов и небольшой размер 2A-последовательностей открывают широкие перспективы использования в биотехнологии векторов на их основе [11]. Однако векторам на основе 2A-последовательностей свойственны и некоторые недостатки. Процессинг 2A-пептидов происходит не со стопроцентной эффективностью, поэтому в клетках образуется небольшое количество нерасщепленной фракции двух белков [15], кроме того в некоторых случаях отмечено значительное увеличение количества нерасщепленной фракции при внедрении в последовательность белка сигнала транспорта в ЭПР [16, 17]. Эффективность расщепления повышается при использовании линкерной последовательности Gly-Ser-Gly или Ser-Gly-Ser-Gly на *N*-конце 2A-пептида. Внедрение такого линкера позволяет уменьшить влияние последовательности предшествующего белка на процессинг 2A-пептида [14, 18].

В настоящее время опубликовано несколько работ, в которых проводилось сравнение векторов на основе *IRES EMCV* и 2A-последовательностей. Так, вектор на основе 2A-последовательности вируса *Thossea asigna* (*T2A*), кодирующий рекомбиназы Flp и Cre, обеспечивал более эффективную сайт-специфическую рекомбинацию по сравнению с вектором на основе *IRES EMCV* [19]. В работе [20] показано, что вектор на основе 2A-последовательности вируса ящура (*F2A*) обеспечивал больший уровень экспрессии *egfp* по сравнению с вектором на основе *IRES EMCV*, однако при этом обнаруживалась низкая эффективность расщепления *F2A*. В приведенных работах использовались неоптимальные варианты 2A-пептида, что могло внести вклад в низкую эффективность расщепления и снизить уровень продукции белков.

Цель данного исследования — сравнить эффективность экспрессии обоих цистронов при использовании бицистронных векторов на основе различных вариантов *IRES EMCV* и *P2A*-последовательности, то есть оценить, какой из векторов обеспечит продукцию наибольшего количества обоих кодируемых белков.

На основе двух вариантов *IRES EMCV* и *P2A*-последовательности нами были созданы бицистронные конструкции, кодирующие флуоресцентные белки EGFP и RFP. Было проведено сравнение уровней продукции репортерных белков, образующихся при транзientной трансфекции трех различных клеточных линий млекопитающих. Уровень экспрессии трансгенов оценивался при помощи флуоресцентной микроскопии, флуориметрии и вестерн-блот-анализа. Во всех клеточных линиях вектор на основе *P2A*-последовательности обеспечивал больший суммарный уровень экспрессии генов *egfp* и *rfp* при высокой эффективности расщепления 2A-пептида.



**Рис. 2.** Особенности использованных в работе генетических конструкций. (а) – Схемы конструкций: R – pCMV-RFP; E – pCMV-EGFP; I – pCMV-RFP-IRESatt-EGFP; I2 – pCMV-RFP-IRES2-EGFP; 2A – pCMV-RFP-P2A-EGFP. *CMV* – промотор цитомегаловируса; *rfp* – последовательность, кодирующая RFP; *egfp* – последовательность, кодирующая EGFP; *IRESatt* – ослабленный вариант *IRES*; *IRES2* – вариант *IRES*, близкий к дикому типу; *P2A* – последовательность, кодирующая P2A-пептид; *PolyA* – сигнал полиаденилирования. (б) – Схема вторичной структуры *IRES EMCV*. AUG – инициаторный кодон, ППТ – полипиримидиновый тракт. (в) – Варианты нуклеотидной последовательности 3'-концевого участка *IRES* и прилегающей к нему нуклеотидной последовательности. *IRES EMCV* – 3'-концевая последовательность *IRES* дикого типа; *IRES2* и *IRESatt* – 3'-концевые последовательности, использовавшиеся в работе. Жирным шрифтом обозначены нуклеотидные остатки, входящие в состав *IRES*. Расположение инициаторного кодона отмечено прямоугольником, подчеркнутым шрифтом обозначен первый кодон *egfp*. (г) – Различные варианты P2A-последовательностей. *P2A w.t.* – нуклеотидная последовательность, кодирующая P2A-пептид дикого типа; *P2A* – нуклеотидная последовательность, использовавшаяся в работе. В нижней части рисунка приведена аминокислотная последовательность P2A-пептида. S(P) – введенная аминокислотная замена. Стрелкой указан сайт расщепления P2A-пептида.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Дизайн бицистронных конструкций на основе *IRES* и *P2A*-последовательности. Для оценки эффективности коэкспрессии трансгенов были созданы три бицистронные конструкции, содержащие репортерные гены флуоресцентных белков RFP и EGFP (рис. 2а). Все конструкции были созданы на базе вектора pGL3, содержали ранний промотор цитомегаловируса (*CMV*) и сигнал полиаденилирования вируса SV40 (*PolyA*). Экспрессия генов осуществлялась под контролем промотора *CMV*. Ген *rfp* располагался ближе к промоторной области, а ген *egfp* после последовательности линкера. В качестве линкера между цистронами две конструкции содержали различные варианты *IRES* вируса EMCV: *IRES2*, последовательность которого была сходна с последовательностью

*IRES* дикого типа, и укороченный *IRES* (*IRESatt*). Третья конструкция в качестве линкера несла *P2A*-последовательность, кодирующую 2A-пептид тешовируса свиней. Также были созданы моноцистронные конструкции, содержащие ген *rfp* или *egfp*.

*IRES EMCV* имеет сложную вторичную структуру (рис. 2б). Последовательности 3'-концов исследуемых в работе *IRES* (*IRESatt* и *IRES2*), а также *IRES* дикого типа представлены на рис. 2в. Укороченный вариант *IRES* имел длину 494 п.о., был лишен нативного инициаторного кодона и связан с последовательностью, кодирующей EGFP, через линкер длиной 22 п.о., последовательность которого отличалась от последовательности *IRES EMCV*. Известно, что удаление нативного инициаторного кодона, а также наличие дополнительных

Таблица 1. Значения флуоресценции RFP и EGFP в клеточных линиях HEK293, M3 и HT1080\*

Клеточная линия	HEK293		M3		HT1080	
	RFP	EGFP	RFP	EGFP	RFP	EGFP
I	27.2+/-6.6	0.8+/-0.4	13.4+/-13.3	0.2+/-0.1	36.1+/- 4.7	0.5+/-0.1
I2	44.7+/-11.4	20.3+/-2.3	36.7+/-12.8	13.5+/-1.6	38.5+/-5.3	7.2+/-0.6
2A	56.5+/-15.8	62+/-10.8	55.8+/-11.7	33.5+/-1.8	34.9+/-3.5	30.7+/-3.1

\* Значения флуоресценции представлены в процентах относительно значений флуоресценции клеток, трансфицированных моноцистронными конструкциями. Обозначения: RFP и EGFP – значения флуоресценции соответствующих белков, I – pCMV-RFP-IRESatt-EGFP; I2 – pCMV-RFP-IRES2-EGFP; 2A – pCMV-RFP-P2A-EGFP. Значения представлены в виде среднего по нескольким экспериментам со стандартной ошибкой среднего (mean+/-SEM, %).

нуклеотидов между РС и последовательностью *IRES* приводит к подавлению активности *IRES* [5, 6], поэтому *IRESatt* предположительно должен проявлять низкую способность к инициации трансляции. *IRES2* имел длину 585 п.о. и был идентичен *IRES* дикого типа за исключением наличия дополнительного остатка аденозина в А6-петле – небольшом структурном элементе, находящимся вблизи 3'-концевого отдела *IRES*. Данная модификация незначительно ослабляет способность *IRES* инициировать трансляцию [5]. *IRES2* содержал нативный инициаторный кодон, за которым были расположены три дополнительных кодона, взятых из вирусной последовательности, что, как показано [8], усиливает активность *IRES*. Поэтому такой вариант *IRES* предположительно должен проявлять более высокую инициаторную активность по сравнению с *IRESatt*.

Последовательность, кодирующая P2A-пептид, соответствовала таковой в геноме тешовируса свиней серотипа 1 [21], однако была подвергнута некоторым модификациям (рис. 2г). Во-первых, отдельные кодоны заменили синонимичными для снижения структурированности нуклеотидной последовательности. Во-вторых, триплет N-концевого остатка Pro был заменен на триплет остатка Ser, что позволило создать линкерную последовательность Gly-Ser-Gly, которая, по литературным данным, увеличивает эффективность расщепления [14, 18]. Кроме того, был добавлен кодон, кодирующий Thr на N-конце, что привело к возникновению сайта узнавания рестриктазой AgeI. Длина P2A-пептида, использовавшегося в работе, составляла 23 а.о.

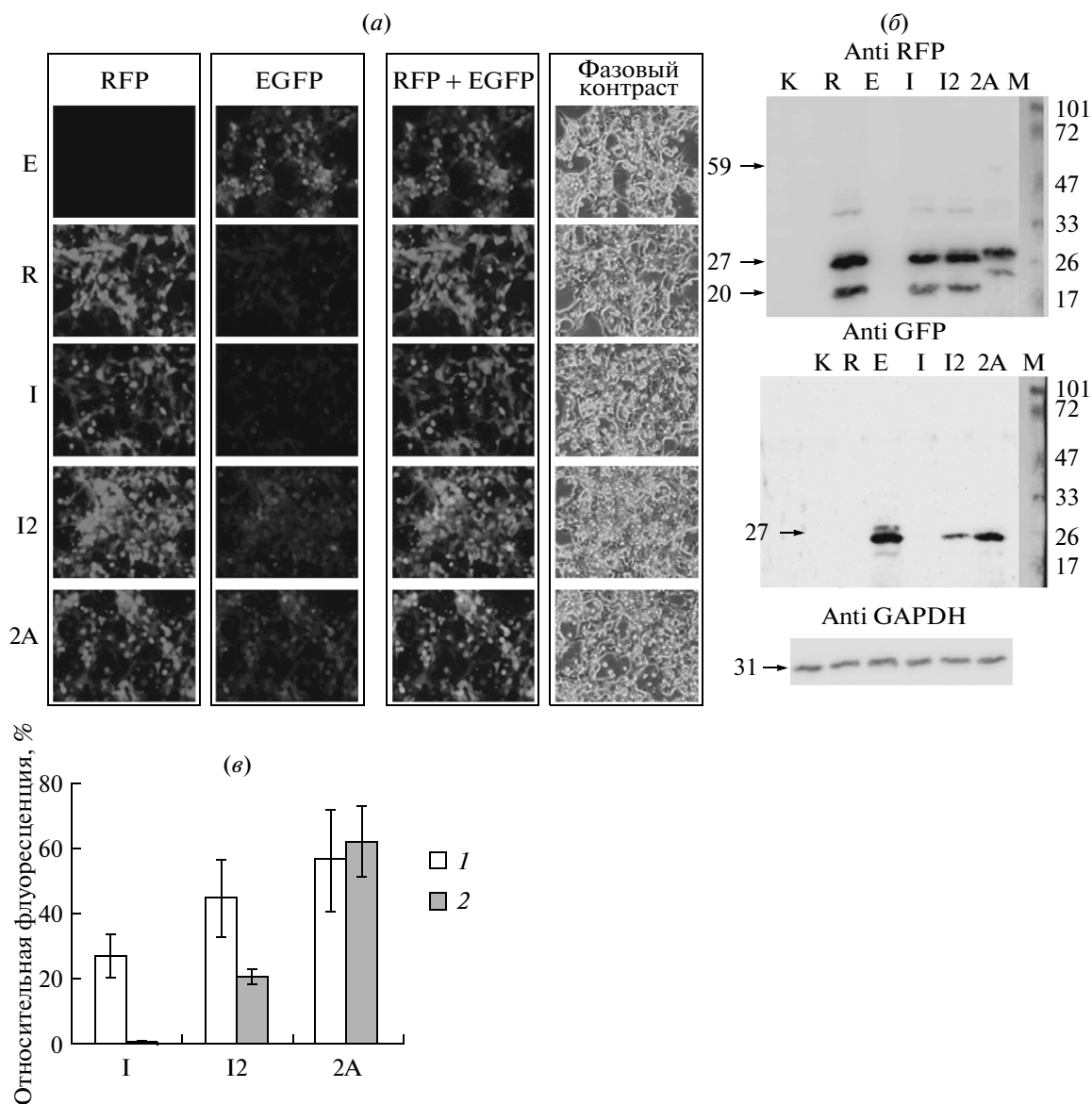
*Изучение эффективности работы полученных конструкций в клетках млекопитающих.* Активность полученных генетических конструкций изучалась в условиях транзиторной трансфекции клеток различных линий млекопитающих. Активность последовательностей *IRES* и *P2A* при использовании векторов в клетках разных типов может варьировать, поэтому эксперименты были проведены на трех хорошо известных клеточных

линиях различного происхождения: HEK293 (почка человеческого эмбриона), M3 (меланома мышцы) и HT1080 (фибросаркома человека). Экспрессию репортерных генов в клетках анализировали путем детекции белков RFP и EGFP с помощью флуоресцентной микроскопии, флуориметрии и вестерн-блота с антителами к белкам RFP и EGFP. Проведение вестерн-блота также позволило оценить степень котрансляционного расщепления RFP-P2A-EGFP. Данные по трем клеточным линиям представлены на рис. 3–5 и в табл. 1.

Согласно результатам флуоресцентной микроскопии, все трансфицированные клетки активно накапливали флуоресцентные белки. В клетках, трансфицированных векторами на основе *IRES2* и *P2A*-последовательности, наблюдался сопоставимый уровень флуоресценции обоих белков, в то время как в клетках, трансфицированных вектором с *IRESatt*, практически не обнаруживалась флуоресценция EGFP (рис. 3а, 4а, 5а).

Флуориметрия позволяет оценивать уровень флуоресценции, а следовательно и содержание репортерных белков RFP и EGFP, на количественном уровне. Значения флуоресценции клеток, трансфицированных бицистронными конструкциями, рассчитывали в процентах от уровня флуоресценции клеток, трансфицированных моноцистронными конструкциями. Уровень флуоресценции последних был принят за 100%. Средние значения флуоресценции представлены в табл. 1 и на рис. 3в, 4в, 5в.

В клеточных линиях HEK293 и M3 вектор на основе *P2A*-последовательности обеспечивал несколько более высокий уровень экспрессии первого цистрона, а вектор с укороченным вариантом *IRES* (*IRESatt*) – наименьший уровень. Данные различия были особенно выражены в линии клеток M3, в которой вектор на основе *IRESatt* обеспечивал наименьший среди всех клеточных линий относительный уровень продукции RFP. Вектор на основе *IRES2* обеспечивал уровень флуоресценции RFP, сопоставимый с уровнем флуоресценции RFP в клетках, трансфицирован-

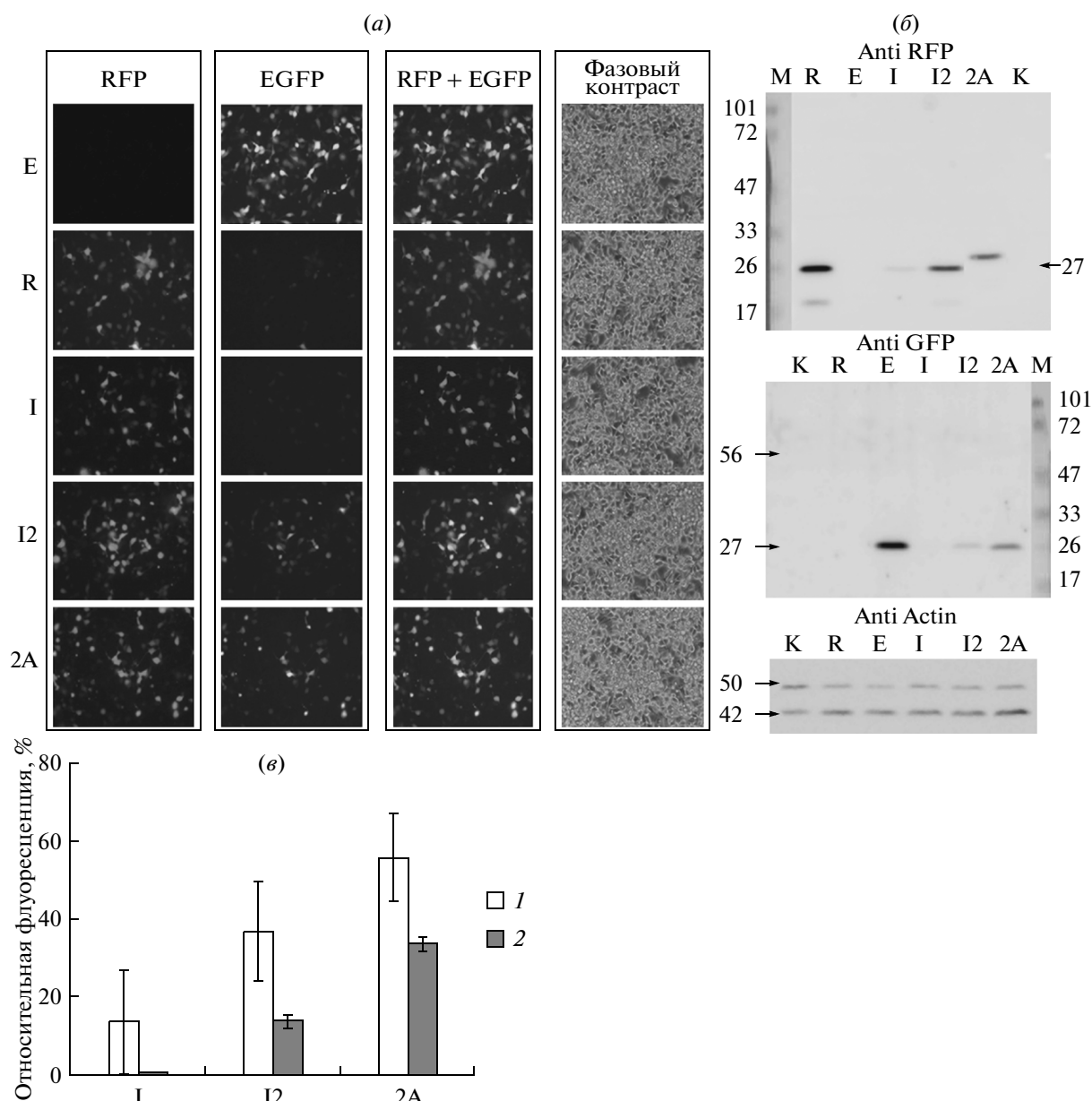


**Рис. 3.** Результаты анализа эффективности работы бицистронных конструкций в клеточной линии HEK293. (а) – Микроскопия клеток HEK293 (X20). В верхней части рисунка обозначены способы получения изображений: RFP и EGFP – микрофотографии с детекцией флуоресценции RFP и EGFP; RFP+EGFP – совмещенное изображение; фазовый контраст – фазово-контрастные микрофотографии при обычном освещении. Слева приведены обозначения векторов, использованных для трансфекции (расшифровку обозначений см. подпись к рис. 2). (б) – Вестерн-блот клеточных лизатов HEK293. AntiRFP и AntiGFP – окраска антителами на RFP и на GFP. Сверху приведены обозначения векторов, использованных для трансфекции (см. рис. 2), К – лизат нетрансфицированных клеток, М – маркеры мол. масс (17–101 кДа). (в) – Флуориметрия клеточной линии HEK293: приведены значения флуоресценции в канале детекции RFP (1) и EGFP (2). Под столбиками приведены обозначения векторов, использовавшихся для трансфекции. Барями обозначена стандартная ошибка среднего по трем независимым экспериментам.

ных вектором на основе *P2A*-последовательности. В клеточной линии HT1080 уровень экспрессии первого цистрона незначительно различался при использовании разных бицистронных векторов и составлял примерно 35% от уровня экспрессии гена *rfp* в клетках, трансфицированных моноцистронной конструкцией.

Во всех клеточных линиях экспрессия второго цистрона *egfp* обнаруживалась на более высоком

уровне при использовании вектора на основе *P2A* по сравнению с векторами на основе *IRE5*. Так, во всех клеточных линиях флуоресценция EGFP была примерно в 3 раза больше в клетках, трансфицированных вектором на основе *P2A*, по сравнению с клетками, трансфицированными *IRE52*. Вектор с *IRE5att* обеспечивал очень низкий уровень экспрессии второго цистрона, разница в уровне флуоресценции EGFP при трансфекции



**Рис. 4.** Результаты анализа эффективности работы бицистронных конструкций в клеточной линии М3. Обозначения указаны на рис. 3.

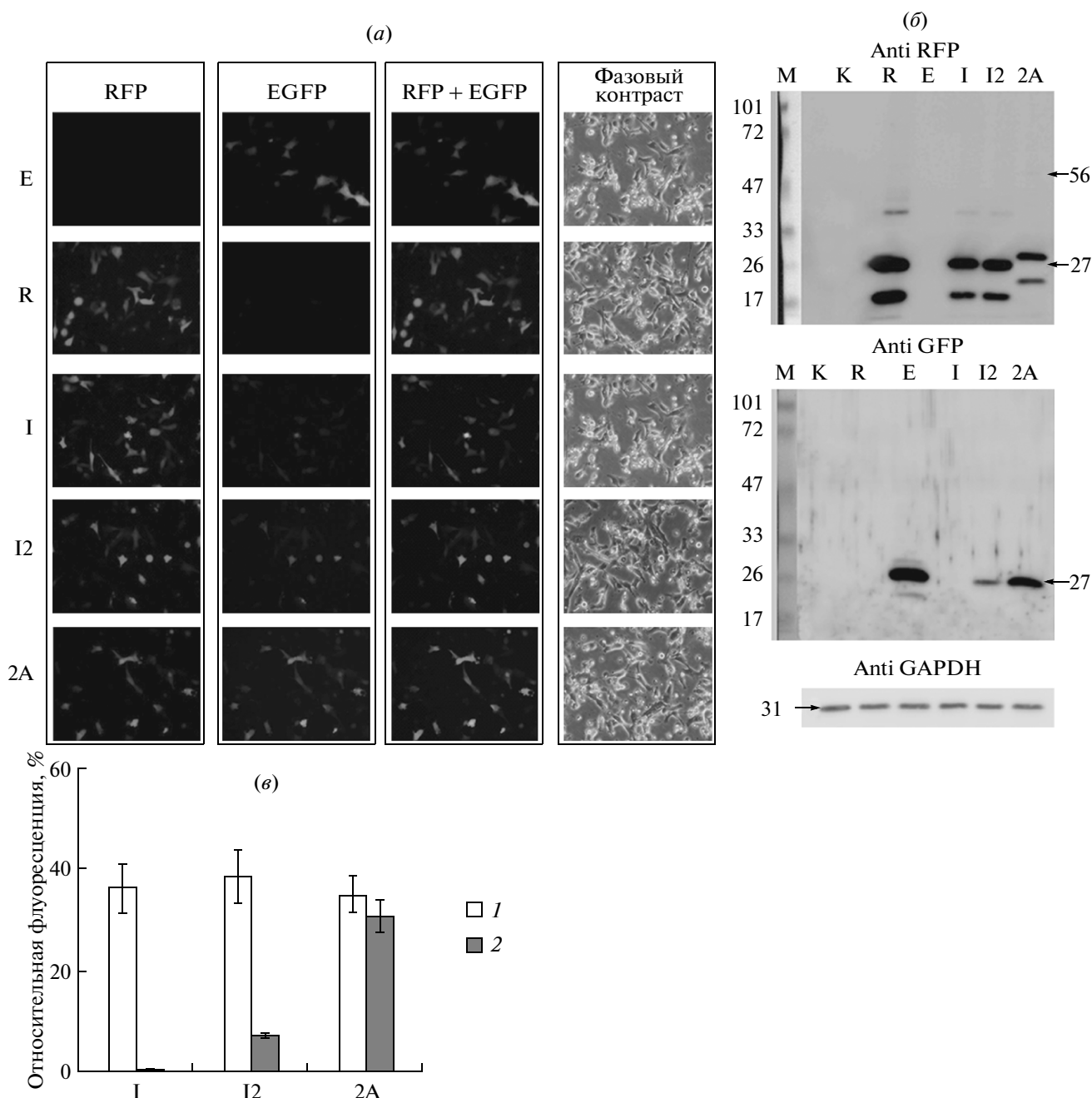
клеток векторами на основе *IRES2* и *IRESatt* составляла 70–170 раз. Таким образом, бицистронный вектор на основе последовательности, кодирующей P2A-пептид, обеспечивал наиболее эффективную экспрессию второго цистрона по сравнению с векторами на основе *IRES*.

Также следует отметить, что во всех клетках, трансфицированных бицистронными конструкциями, уровень экспрессии репортерных генов был ниже, чем в клетках, трансфицированных моноцистронными векторами.

Результаты вестерн-блота подтвердили данные флуориметрии (рис. 3б, 4б, 5б). В случае вектора на основе *IRESatt* детектировать EGFP с по-

мощью вестерн-блота не удалось: его содержание, по всей видимости, находится за пределами чувствительности метода.

Расчетная масса белков EGFP и RFP составляет 27 кДа. В лизатах клеток, трансфицированных вектором на основе последовательности P2A, окрашенных антителами на RFP, наблюдалось смещение основной полосы в связи с тем, что гибридный белок RFP-P2A несет дополнительные 22 а.о. на C-конце, за счет которых молекулярная масса белка увеличилась до 29 кДа. При вестерн-блоте кроме основной полосы белка RFP с M27 кДа обнаруживались дополнительные полосы, соответствующие белкам с массой около 20 и 40 кДа.



**Рис. 5.** Результаты анализа эффективности работы бицистронных конструкций в клеточной линии HT1080. Обозначения как на рис. 3.

Хорошо известно, что многие флуоресцентные белки способны формировать гомодимеры, особенно в концентрированных растворах, поэтому полосе с массой около 40 кДа может соответствовать продукт димеризации RFP или продукт димеризации фрагментов его расщепления. Появление полосы с массой около 20 кДа можно объяснить расщеплением белка RFP в области хромофорной группы, которое происходило при кипячении клеточных лизатов перед проведением электрофореза. Такое расщепление белковой

цепи при кипячении описано для красного флуоресцентного белка DsRed, при этом авторы работы предположили, что расщеплению подвержены только те флуоресцентные белки, хромофоры которых содержат ацилиминогруппу [22]. Как было выяснено при анализе кристаллической структуры, хромофор красного флуоресцентного белка Tag-RFP также содержит ацилиминогруппу [23]. Для выравнивания количества белка в образцах был проведен вестерн-блот тех же образцов с ан-

**Таблица 2.** Относительное содержание нерасщепленного белка RFP-P2A-EGFP в лизатах клеток, трансфицированных вектором на основе P2A-пептида

Клеточная линия	НЕК293	М3	HT1080
Доля нерасщепленной фракции, % от RFP	1.6	Не обнаружена	1.3

тителами к белку домашнего хозяйства GAPDH или к актину.

Содержание нерасщепленного белка RFP-P2A-EGFP в лизатах оценивалось по интенсивности люминесценции полосы, соответствующей нерасщепленному белку с массой 56 кДа. Значения люминесценции данной полосы рассчитывали относительно суммарной люминесценции всех полос, соответствующих RFP в дорожке P2A-пептида.

Во всех клеточных линиях наблюдалась высокая эффективность расщепления P2A-пептида (табл. 2). Максимальный выход химерного белка наблюдался в клеточной линии НЕК293, в которой его количество достигало 1.6% от общего количества RFP. В клеточной линии М3 присутствие нерасщепленного белка детектировать не удалось.

Полученные данные убедительно показывают, что бицистронный вектор на основе P2A-последовательности более эффективен, чем векторы на основе IRES, так как при его использовании удается достичь большей суммарной экспрессии обоих цистронов. При этом в клетках, трансфицированных вектором на основе P2A-последовательности, наблюдается накопление нерасщепленной фракции белков, связанных через 2A-пептид, но их количество незначительно (меньше 2%) и не может являться существенным недостатком использования вектора данного типа.

Наши результаты, полученные при оценке степени расщепления P2A-пептида в клеточной линии HT1080, расходятся с данными работы [15]. В этой работе степень расщепления P2A-пептида в клеточной линии HT1080 равнялась 80–85%, в то время как в нашей работе эта величина составляла около 99%. Такое расхождение может быть связано с различиями в белковом окружении P2A-пептида, которое способно влиять на эффективность его расщепления. В нашей работе P2A связывал последовательности RFP и EGFP (RFP-P2A-EGFP), в то время как в работе [15] P2A-пептид связывал последовательности EGFP и mCherry (EGFP-P2A-mCherry). Кроме того, использованный в нашей работе вариант P2A отличался наличием остатка Thr на N-конце. В нашем случае нуклеотидная последовательность, кодирующая P2A, отличалась по кодонному составу от последовательности, использованной в работе [15], что также могло внести вклад в обнаруженные различия.

Повышенный уровень экспрессии цистронов при использовании вектора на основе P2A можно объяснить как большей скоростью синтеза репортерных белков, так и пониженной скоростью деградации мРНК и белков. При трансляции мРНК *rfp-P2A-egfp* оба целевых белка образуются в результате кэп-зависимой инициации, в то время как в случае с IRES-содержащей мРНК продукт второго цистрона образуется в результате внутренней инициации трансляции, которая может быть менее эффективна по сравнению с кэп-зависимой инициацией, особенно при использовании ослабленных вариантов IRES [9]. Обнаруженные различия также могут быть объяснены ускоренной деградацией мРНК, содержащих IRES, так как такие молекулы мРНК подвержены распознаванию системой NMD [10]. Кроме того, присутствие на N-конце EGFP остатка Pro, остающегося при процессинге P2A, может уменьшать скорость деградации белка [24], что ведет к увеличению его количества в клетке.

Интересным является сравнение эффективности экспрессии цистронов в клетках, трансфицированных вектором на основе IRESatt и IRES2. В использованных клеточных линиях экспрессия обоих цистронов была ниже при трансфекции вектором pCMV-RFP-IRESatt-EGFP. IRESatt инициировал трансляцию цистрона *egfp* значительно менее эффективно по сравнению с IRES2. Так, уровни экспрессии второго цистрона при трансфекции векторами на основе IRES2 и IRESatt различались более чем на порядок. При этом разница в уровне экспрессии *rfp* не была столь значительной.

Больше всего (примерно в три раза) уровень экспрессии первого цистрона при трансфекции клеток двумя бицистронными векторами на основе IRES2 и IRESatt различался в клеточной линии М3. Такой результат можно объяснить повышенной деградацией образующейся мРНК, содержащей IRESatt. Возможно, активная трансляция второго цистрона предохраняет мРНК с IRES2 от деградации, вызванной присутствием в середине молекулы терминирующего кодона, который может распознаваться системой деградации NMD как преждевременный терминирующий кодон [10, 25]. По последним данным, для деградации мРНК с преждевременным терминирующим кодоном необходимо взаимодействие белка UPF1 (Up-frameshift suppressor 1 homolog) с терминирующими рибосомами и последующее движение



Таблица 3. Последовательности праймеров, использовавшихся в работе

Название праймера	Нуклеотидная последовательность, 5'-3'
RFP for	GAATTCTGCAGTCGACGGTAC
RFP rev	AGAGAAGTTCGTGGCTCCGGAACCGGTATTAAGTTTGTGCCCCAG
EGFP for	AGACGTGGAGGAGAATCCCGGGCCTATGGTGAGCAAGGGCGAG
EGFP rev	ATATCTAGATTACTTGTACAGCTCGTCCATG
2A for	AGCAGGAGACGTGGAGGAGAATCCCG
2A rev	TGCTTCAGCAGAGAGAAGTTCGTGGCTCCG
IRES for	GTTATTTTCCACCATATTGCCGTC
IRES rev	GATCGTGTTTTTCAAAGGAAAACC

его к 3'-концу мРНК [26]. Можно предположить, что рибосомы, осуществляющие трансляцию второго цистрона, препятствуют движению UPF1 к 3'-концу мРНК. Так как при использовании вектора на основе *IRESatt* трансляция второго цистрона значительно ослаблена, мРНК, содержащая данный элемент, может деградировать быстрее мРНК, содержащей активный *IRES2*.

### ЗАКЛЮЧЕНИЕ

В настоящее время *IRES EMCV* часто применяется для создания генетических векторов. Однако подчас исследователи используют *IRES* с неоптимальной структурой, что приводит к уменьшению продукции целевых белков. Ослабленный вариант *IRES* в бицистронных векторах в нашем исследовании приводил к падению уровня синтеза продуктов обоих цистронов. Для эффективной экспрессии обоих цистронов следует использовать *IRES EMCV*, близкий к *IRES* дикого типа. При этом особое внимание следует уделять 3'-концевой последовательности *IRES* – для максимальной экспрессии следует сохранять последовательность фрагмента ППТ-AUG вирусного *IRES*.

При трансфекции клеток вектором на основе последовательности, кодирующей P2A-пептид, наблюдалась большая суммарная экспрессия обоих цистронов по сравнению с векторами на основе *IRES*, при этом эффективность процессинга пептида 2A была крайне высока, поэтому использование векторов на основе P2A-последовательности предпочтительно в тех биотехнологических задачах, где требуется высокий уровень экспрессии трансгенов и концевые модификации белков несущественны.

### ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

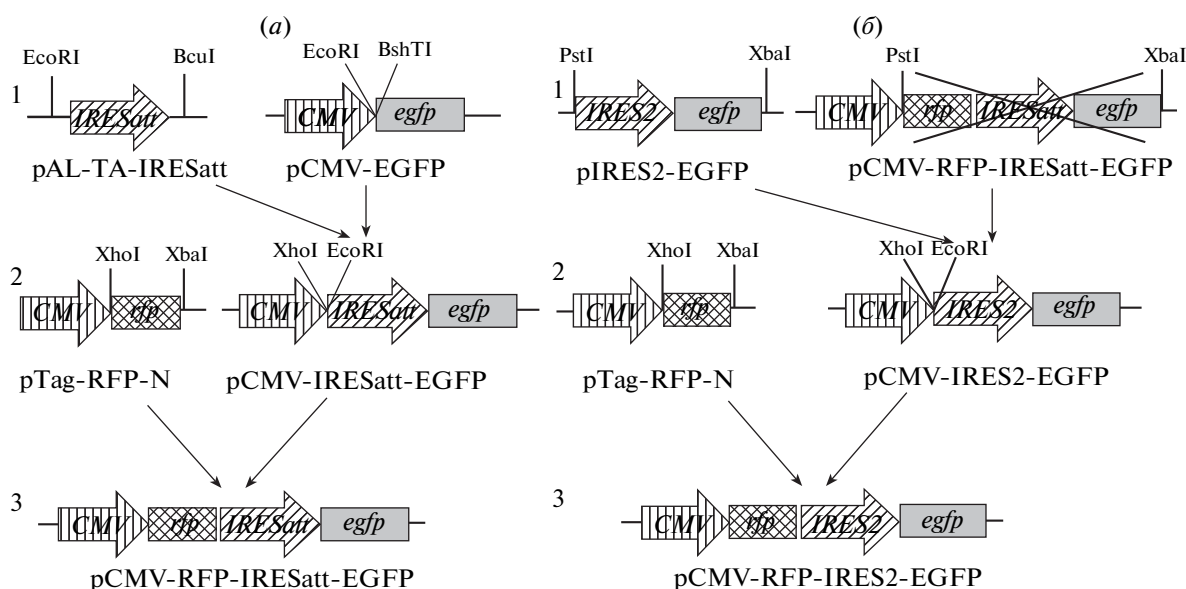
**Культуры клеток и ткани.** В работе использовали клетки линий НЕК293 (почка эмбриона человека, CRL-1573, ATCC), М3 (меланома мыши, CCL-53.1, ATCC), НТ1080 (фибросаркома человека, CCL-121, ATCC). Клетки культивировали

в среде DMEM/F12 (1 : 1), содержащей 10% эмбриональной телячьей сыворотки, 100 ед./мл пенициллина, 100 мкг/мл стрептомицина и 0.25 мкг/мл амфотерицина (Invitrogen, США), при 37°C в атмосфере с 5% CO<sub>2</sub>.

**Создание генно-инженерных конструкций.** Все исследуемые в работе конструкции были созданы на основе pGL3 basic-вектора (Clontech, США), из которого был вырезан цистрон люциферазы светлячка, но клонирован промотор *CMV*. При создании векторов pCMV-RFP и pCMV-EGFP цистроны флуоресцентных белков были получены из плазмид pEGFP-N1 (Clontech, США) и pTag-RFP-N (Евроген, Россия) путем расщепления по сайтам XbaI/HindIII (при создании pCMV-EGFP) или XbaI/XhoI (при создании pCMV-RFP). По тем же сайтам проводилось лигирование полученных цистронов в pGL3-вектор.

Создание вектора pCMV-RFP-IRESatt-EGFP проходило в несколько этапов. На первом этапе последовательность *IRES* из плазмиды pFB-Neo (Stratagene, США) была амплифицирована при помощи праймеров IRES for и IRES rev (табл. 3), затем амплифицированный фрагмент был лигирован в pAL-TA-вектор (Евроген, Россия). На втором этапе последовательность, кодирующая *IRESatt*, была вырезана из плазмиды pAL-TA-IRESatt по сайту BcuI с последующей достройкой 3'-концов, а также по сайту EcoRI (рис. 6a, 1), после чего клонирована в вектор pCMV-EGFP, который был предварительно гидролизован по сайту BshT1 с достройкой 3'-концов, а также по сайту EcoRI (рис. 6a, 1). В результате была получена плаزمид pCMV-IRESatt-EGFP (рис. 6a, 2). Далее цистрон *rfp* был вырезан из плазмиды pTag-RFP-N путем гидролиза по сайту XbaI с последующей достройкой 3'-концов, а также по сайту XhoI (рис. 6a, 2), после чего клонирован в вектор pCMV-IRESatt-EGFP, который был гидролизован по сайту EcoRI с достройкой 3'-концов и по сайту XhoI. В результате была получена плазмид pCMV-RFP-IRESatt-EGFP (рис. 6a, 3).

Для создания вектора pCMV-RFP-IRES2-EGFP на первом этапе кассета *IRES2-egfp* была вырезана из вектора pIRES2-EGFP (Clontech, США) пу-



**Рис. 6.** Схемы получения векторов на основе *IRES*. (а) — Получение вектора pCMV-RFP-IRESatt-EGFP (б) — получение вектора pCMV-RFP-IRES2-EGFP. Под схемами векторных конструкций отображены их названия, на схемах указаны входящие в состав векторов элементы, а также сайты рестрикции, по которым осуществлялся гидролиз. *CMV* — промотор цитомегаловируса; *rfp* — последовательность, кодирующая RFP; *egfp* — последовательность, кодирующая EGFP; *IRESatt* — ослабленный вариант *IRES*; *IRES2* — вариант *IRES*, близкий к дикому типу. Слева цифрами 1–3 указаны этапы клонирования.

тем гидролиза по сайтам XbaI и PstI (рис. 6б, 1), после чего клонирована в вектор pCMV-RFP-IRESatt-EGFP, из которого путем гидролиза по сайтам XbaI и PstI был вырезан участок *rfp-IRESatt-egfp*. В результате была получена плаزمида pCMV-IRES2-EGFP (рис. 6б, 2). На следующем этапе цистрон *rfp* был вырезан из плазмиды pTag-RFP-N путем гидролиза по сайту XbaI с последующей достройкой 3'-концов, а также по сайту XhoI (рис. 6б, 2), после чего клонирована в вектор pCMV-IRES2-EGFP, который был гидролизован по сайту EcoRI с последующей достройкой 3'-концов и по сайту XhoI. В результате была получена плазмида pCMV-RFP-IRES2-EGFP (рис. 6б, 3).

Последовательность *rfp-P2A-egfp* получали при помощи ПЦР в два этапа (рис. 7). На первом этапе на матрице плазмиды pEGFP-N1 при помощи праймеров EGFP for/EGFP rev амплифицировали последовательность *egfp*, на матрице плазмиды pTag-RFP-N при помощи праймеров RFP for/RFP rev — последовательность *rfp*. Праймеры RFP rev и EGFP for несли фрагменты *P2A*-последовательности. Для амплификации использовалась смесь ДНК-полимераз Encuslo (Евроген, Россия). После окончания ПЦР продукт был очищен на колонке Wizard SV Gel and PCR Clean-Up System (Promega, США). На втором этапе амплификации использовались следующие пары праймеров: RFP for/2A rev в одной реакционной смеси и 2A for/EGFP rev в другой. Праймеры 2A rev и 2A for были комплементарны нуклеотидной последовательности, введенной на первом этапе ПЦР.

Праймеры 2A rev и 2A for предварительно были подвергнуты реакции фосфорилирования Т4-полинуклеотидкиназой (Fermentas, Литва). Для амплификации была использована Pfu-ДНК-полимераза (Fermentas, Литва), которая, в отличие от смеси полимераз Encuslo, не присоединяет на 3'-конец синтезируемой цепи выступающий остаток аденина. Продукты реакции очищали на колонках Wizard SV Gel and PCR Clean-Up System (Promega, США). На третьем этапе полученные ПЦР-продукты были гидролизованы по сайтам PstI и XbaI, содержащимся в последовательности праймеров RFP for и EGFP rev, и клонированы в вектор pCMV-RFP-IRES2-EGFP, который был предварительно гидролизован по сайтам PstI и XbaI (это позволило вырезать участок *rfp-IRES2-egfp*). В результате была получена генетическая конструкция pCMV-RFP-P2A-EGFP.

Полученные плазмиды нарабатывали в штамме DH5α *E. coli* и выделяли в количестве 100–150 мкг с помощью набора QIAGEN Plasmid Midi Kit (Qiagen, Германия).

*Трансфекцию клеток* проводили в 6-луночных планшетах липофектаминоом 2000 (Invitrogen, США) согласно рекомендациям производителя, используя 4 мкг плазмидной ДНК на одну лунку.

После трансфекции клетки культивировали при 37°C 24 ч в среде DMEM/F12 (1 : 1) без антибиотиков, содержащей 10% эмбриональной телячьей сыворотки, затем среду сменяли на среду с антибиотиком и инкубировали при 37°C 24 ч.

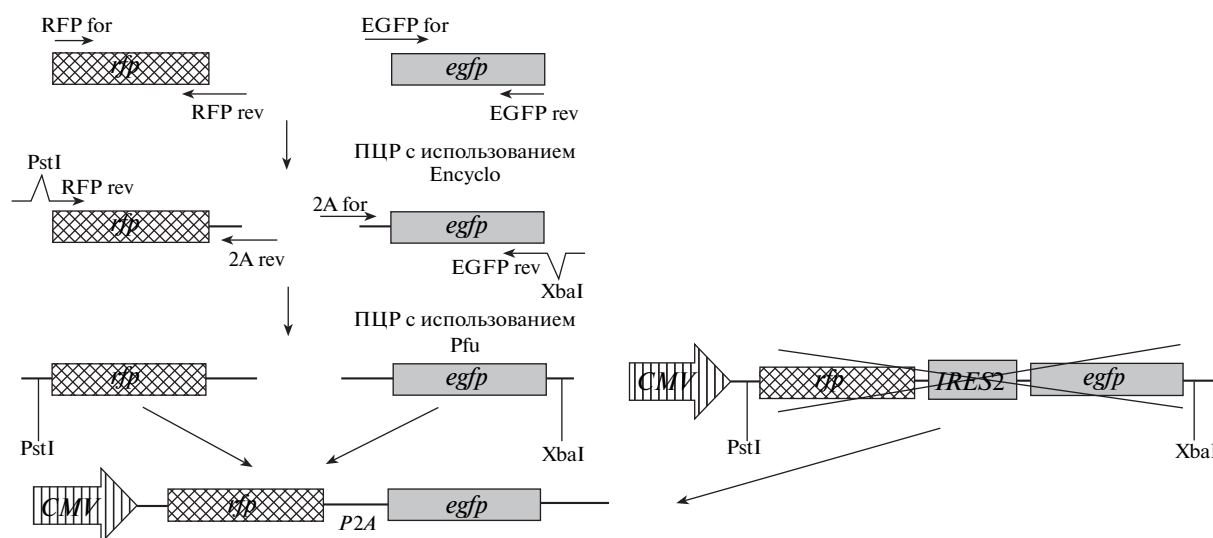


Рис. 7. Схема получения вектора pCMV-RFP-P2A-EGFP. Обозначения как на рис. 2–6. Также на схеме приведены использованные праймеры и рестриктазы.

Флуоресцентную микроскопию клеток проводили через 48 ч после трансфекции при помощи фазово-контрастного микроскопа Nikon Eclipse TE2000-U (Nikon, Нидерланды). Возбуждение флуоресценции EGFP осуществляли при помощи светофильтра В-2А, возбуждение флуоресценции RFP – при помощи светофильтра G2-А. Также были получены фазово-контрастные микрофотографии. При помощи программы Adobe Photoshop CS3 проводили совмещение полученных флуоресцентных изображений. Все микрофотографии были сделаны при помощи окуляра, дающего увеличение в 20 раз.

**Измерение красной и зеленой флуоресценции клеток.** Через 48 ч после трансфекции суспензию с суммарным количеством 250 000 клеток переносили в черный 96-луночный планшет и доводили до объема 200 мкл буфером PBS (концентрацию клеток измеряли при помощи прибора Cell Counter TC 10 (Bio-Rad Laboratories, США).

В работе определяли флуоресценцию двух белков: EGFP и RFP. Измерения проводили на микропланшетном флуориметре “GENios Pro” (Tecan, Австрия). При измерении флуоресценции EGFP использовали светофильтр 485 нм (+/-20) для возбуждения флуоресценции и светофильтр 535 нм (+/-25) для детекции флуоресценции. При измерении флуоресценции RFP использовали пару светофильтров 535 нм (+/-25)/590 нм (+/-10). Измерение флуоресценции для каждой серии трансфицированных клеток проводили дважды. Трансфекцию каждой клеточной линии проводили не менее трех раз. Рассчитывали среднее значение флуоресценции каждого белка, для каждой клеточной линии определяли стандартную ошибку среднего.

**Вестерн-блот.** Собранные клеточные осадки растворяли в деионизованной воде, а затем лизировали равным объемом буфера 2× Sample (Tris-Glycine SDS Sample Buffer Novex (Invitrogen, США) с 4% β-меркаптоэтанолом) из расчета конечной концентрации клеток 1000 клеток на мкл. Лизаты белков разделяли электрофорезом в 12% полиакриламидном геле в присутствии додецилсульфата натрия и переносили на мембрану PVDF Immobilone-P (Millipore, США). После блокировки 5% сухим молоком мембрану обрабатывали первичными антителами к репортерным белкам EGFP и RFP, а также к белкам домашнего хозяйства GAPDH или актину. В работе использовали следующие разведения антител: кроличьи антитела anti-RFP 1 : 33 333 (Евроген, Россия); мышинные антитела anti-EGFP 1 : 1000, anti-GAPDH 1 : 5000, anti-actin 1 : 400 (Santa Cruz Biotechnology, США). Мембрану отмывали 4 раза по 15 мин раствором PBS-T, а затем инкубировали с вторичными антителами к иммуноглобулинам мыши или кролика в разведении 1 : 1000 (Santa Cruz Biotechnology, США), конъюгированными с пероксидазой хрена, и снова отмывали буфером PBS-T. Мембрану проявляли при помощи системы реагентов Immun-Star HRP Chemiluminescent (Bio-Rad Laboratories, США). Люминесценцию регистрировали с помощью прибора Bio-Rad VersaDoc MP4000 (Bio-Rad Laboratories, США), данные обрабатывали, используя компьютерную программу Quantity One. Для каждой мембраны делали два снимка: на одном снимке детектировали люминесценцию, второй делали в видимом свете и на нем детектировали маркеры молекулярных масс. Далее получали совмещенное изображение снимков при помощи программы Adobe Photoshop CS3.

## БЛАГОДАРНОСТИ

Работа поддерживалась программой Президиума РАН “Молекулярная и клеточная биология”, а также программой “Ведущие научные школы РФ” (НШ45638.2010.4).

## СПИСОК ЛИТЕРАТУРЫ

1. *de Felipe P.* // *Curr Gene Ther.* 2002. V. 2. P. 355–378.
2. *Pestova T.V., Kolupaeva V.G., Lomakin I.B., Pilipenko E.V., Shatsky, I.N., Agol V.I., Hellen, C.U.* // *Proc. Natl. Acad. Sci. U.S.A.* 2001. V. 98. P. 7029–7036.
3. *Jang S.K., Kräusslich H.G., Nicklin M.J., Duke G.M., Palmenberg A.C., Wimmer, E.* // *J. Virol.*, 1988. V. 62. P. 2636–2643.
4. *Kaminski A., Howell M.T., Jackson R.J.* // *EMBO J.* 1990. V. 9. P. 3753–3759.
5. *Bochkov Y.A., Palmenberg A.C.* // *BioTechniques.* 2006. V. 41. P. 283–288.
6. *Martin P., Albagli O., Poggi M.C., Boulukos K.E., Pognonec P.* // *BMC Biotechnol.* 2006. V. 6.
7. *Jang S.K., Wimmer E.* // *Genes Dev.* 1990. V. 4. P. 1560–1572.
8. *Qiao J., Roy V., Girard M.H., Caruso M.* // *Hum. Gene Ther.* 2002. V. 13. P. 881–887.
9. *Mizuguchi H., Xu Z., Ishii-Watabe A., Uchida E., Hayakawa T.* // *Mol. Ther.* 2000. V. 1 P. 376–382.
10. *Shikama Y., Hu H., Ohno M., Matsuoka I., Shichishima T., Kimura J.* // *BMC Mol. Biol.* 2010. V. 11.
11. *de Felipe P., Luke G.A., Hughes L.E., Gani D., Halpin C., Ryan M.D.* // *Trends Biotechnol.* 2006. V. 24. P. 68–75.
12. *Agol V.I., Gmyl A.P.* // *Nat. Rev. Microbiol.* 2010. V. 8. P. 867–878.
13. *Donnelly M.L., Hughes L.E., Luke G., Mendoza H., ten Dam E., Gani D., Ryan M.D.* // *J. Gen. Virol.* 2001. V. 82. P. 1027–1041.
14. *Szymczak A.L., Workman C.J., Wang Y., Vignali K.M., Dilioglou S., Vanin E.F., Vignali D.A.* // *Nat. Biotechnol.* 2004. V. 22. P. 589–594.
15. *Kim J.H., Lee S.R., Li L.H., Park H.J., Park J.H., Lee K.Y., Kim M.K., Shin B.A., Choi S.Y.* // *PLoS ONE* 2011. V. 6. P. e18556.
16. *de Felipe P., Luke G.A., Brown J.D., Ryan M.D.* // *Bio-technol J.* 2010. V. 5. P. 213–223.
17. *Rothwell D.G., Crossley R., Bridgeman J.S., Sheard V., Zhang Y., Sharp T.V., Hawkins R.E., Gilham D.E., McKay T.R.* // *Hum. Gene Ther.* 2010. V. 21. P. 1631–1637.
18. *Yang S., Cohen C.J., Peng P.D., Zhao Y., Cassard L., Yu Z., Zheng Z., Jones S., Restifo N.P., Rosenberg S.A., Morgan R.A.* // *Gene Ther.* 2008. V. 15. P. 1411–1423.
19. *Anderson R.P., Voznyanova E., Voznyanov Y.* // *Nucleic Acids Research.* 2012. V. 40 P. 1–9.
20. *Chan H.Y., V.S., Xing X., Kraus P., Yap S.P., Ng P., Lim S.L., Lufkin T.* // *PLoS ONE.* 2011. V. 6. P. e28885.
21. *Doherty M., Todd D., McFerran N., Hoey E.M.* // *J. Gen. Virol.* 1999. V. 80. P. 1929–1941.
22. *Gross L.A., Baird G.S., Hoffman R.C., Baldrige K.K., Tsien R.Y.* // *Proc. Natl. Acad. Sci. U.S.A.* 2000. V. 97. P. 11 990–11 995.
23. *Subach O.M., Malashkevich V.N., Zencheck W.D., Morozova K.S., Piatkevich K.D., Almo S.C., Verkhusha V.V.* // *Chem. Biol.* 2010. V. 17 P. 333–341.
24. *Varshavsky A.* // *Protein Sci.* 2011. P. 1298–1345.
25. *Schwartz A.M., Komarova T.V., Skulachev M.V., Zvereva A.S., Dorokhov I.L., Atabekov J.G.* // *Biochemistry Mosc.* 2006. V. 71. P. 1377–1384.
26. *Shigeoka T., Kato S., Kawaichi M., Ishida Y.* // *Nucleic acids research.* 2012. V. 40. P. 6887–6897.

## Quantitative Comparison of Expression for Genes Linked in Bicistronic Vectors Via Ires or 2A-Peptide of Porcine Teschovirus-1 Sequence

A. I. Kuzmich<sup>#</sup>, A. V. Vvedenskii, E. P. Kopantzev, T. V. Vinogradova

<sup>#</sup> Phone: +7 (495) 330-69-92; fax: +7 (495) 330465438; e-mail: akrubik@gmail.com

*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, 117997, Moscow, GSP-7, Miklukho-Maklaya, 16/10*

Simultaneous expression of multiple target genes is often required in biotechnology. Multicistronic vectors coding for several proteins are being actively developed for this purpose. In commercially available vectors different variants of encephalomyocarditis virus internal ribosome entry site (IRES EMCV) are used most often. However, many researchers consider that utilization of self-cleaving 2A peptides sequences within multi- and bicistronic vectors is more promising. In this work, we compared the efficiency of gene expression in cells transfected with bicistronic vectors based on IRES EMCV and 2A peptide sequence derived from porcine teschovirus-1 (P2A). Efficiency of gene expression was determined in three mammalian cell lines by measurement of co-expression levels of genes coding for RFP and EGFP proteins linked by IRES or P2A sequence. Higher level of transgene expression was exhibited by cells transfected with the vector containing the 2A peptide sequence.

*Keywords: bicistronic vector, IRES, EMCV, 2A peptide, porcine teschovirus*