

## КРАТКИЕ СООБЩЕНИЯ

УДК 577.29

©Курбатов, Згода

### ИСПОЛЬЗОВАНИЕ ДНК-МАТРИЦЫ В КАЧЕСТВЕ КОНТРОЛЯ ПРИ ПРОВЕДЕНИИ ТРАНСКРИПТОМНОГО АНАЛИЗА НА СТАНДАРТНЫХ МИКРОЧИПАХ AGILENT TECHNOLOGIES

Л.К. Курбатов\*, В.Г. Згода

Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича,  
119121, Москва, ул. Погодинская, д. 10, стр.8; тел. +7(499)246-69-80; факс: +7 (499) 245-08-57;  
эл. почта: kurbatovl@mail.ru

Достоверность результатов транскриптомного анализа на микрочипах достигается многоступенчатой нормализацией полученных данных. Одно из важных условий для такой нормализации – наличие внутренних контролей. При подготовке проб по стандартному протоколу фирмы Agilent Technologies в качестве таких контролей используется набор Spike-In Kit, представляющий собой смесь разных РНК в известных соотношениях. Синтезированные *in vitro* молекулы РНК применялись в исследованиях зависимости между интенсивностью флуоресцентного сигнала и наличием РНК в образце. К минусам таких контролей можно отнести сравнительную сложность их получения и нестабильность РНК. Поскольку основой для синтеза меченой кРНК при пробоподготовке по протоколу Agilent является наличие Т7 промотора, мы предположили, что таким же успешным будет синтез любой последовательности, имеющей в начале такой промотор. Более того, в качестве матрицы может выступать молекула ДНК. В ходе нашей работы на примере нескольких генов, включенных в плазмидную ДНК после Т7-промотора, было показано, что такая конструкция является полноценной матрицей для синтеза РНК с включением флуоресцентной метки. По сравнению с классическим типом внутренних контролей на основе РНК, такая матрица отличается стабильностью, простотой изготовления и легко может быть получена в больших количествах.

**Ключевые слова:** микрочипы, плазмидная ДНК, Т7-промотор, уровень флуоресценции, меченая кРНК

**DOI** 10.18097/PBMC20166206715

## ВВЕДЕНИЕ

В процессе совершенствования технологии транскриптомного анализа на микрочипах почти сразу возникла необходимость во внутренних контролях, сигнал от которых в дальнейшем использовался для нормализации и стандартизации данных. Кроме того, некоторые исследователи рассматривали возможность оценки количества мРНК в пробе по величине флуоресцентного сигнала на микрочипе [1-3]. Традиционным подходом в решении этой задачи является использование экзогенных нуклеотидных последовательностей, которые вносятся в исследуемую пробу на каком-то этапе до гибридизации [4-6]. Такими референсными последовательностями могут быть синтезированные олигонуклеотиды [1] или полученные *in vitro* молекулы РНК определённых генов. Второй подход используется чаще и предполагает, как правило, синтез нужных последовательностей с помощью ПЦР на матрице геномной ДНК или кДНК с последующим включением в специальный вектор для экспрессии [4-6].

В исследованиях с применением внутренних контролей при пробоподготовке и гибридизации кРНК/(кДНК) на микрочипах во многих случаях было показано, что такой подход оправдывает себя как метод нормализации и стандартизации данных [7]. В сочетании с ПЦР в реальном времени и соответствующей статистической обработкой сигнал от таких референсных РНК позволяет оценить абсолютное количество матрицы [2] и рассчитать исходное количество копий РНК на клетку [3].

В нашей работе на примере нескольких генов были воспроизведены условия стандартной пробоподготовки по протоколу фирмы “Agilent Technologies” (США) для двухцветной гибридизации, при этом в качестве матрицы использовалась плазмидная ДНК с фрагментами контрольных генов под контролем Т7-промотора. После гибридизации на стандартном микрочипе была произведена оценка зависимости сигнала от количества исходного материала в нескольких разведениях.

## МЕТОДИКА

В качестве матрицы было использовано два типа генетических конструкций, имеющих на 5'-конце генной последовательности Т7 промотор. Первый тип представлял собой короткие фрагменты генов *ROCK1*, *SMAD2* и *SMAD4*, комплементарные последовательностям зондов на микрочипах, полученные с помощью ПЦР. Размер ПЦР-продуктов составлял порядка 130 нуклеотидов. В качестве матрицы использовалась кДНК клеточной линии HepG2. Синтезированные последовательности были клонированы в плазмиду pGEM-T-Easy, предназначенную для лигирования с ПЦР-продуктами. Наличие вставки в плазмидах оценивали по результатам рестрикции и электрофореза в 1% агарозном геле. Плазмиды со вставкой нужного размера были отсекарованы. Второй тип – полноразмерные гены в коммерческих конструкциях для индуцированной экспрессии фирмы “GeneCopoeia” (США). В данном случае это были гены цитохромов

\* - адресат для переписки

*CYP2D6*, *CYP2E1*, *CYP1A2*, а также ген *POR*, клонированные в виде ORF (открытая рамка чтения) в вектор pReceiver-B01. Концентрацию ДНК определяли на приборе NanoDrop2000 ("Thermo Scientific", США). Пробы для анализа представляли собой смесь плазмид в расчетной концентрации, соответствующей 6 выбранным точкам с условным содержанием от 1 до 1000 молекул на одну клетку образца. Пробоподготовку осуществляли в соответствии с протоколом Two Color Microarray-Based gene Expression Analysis Version 5.5 для двухцветной гибридизации и использованием набора Agilent Low RNA Input Fluorescent Linear Amplification Kit ("Agilent Technologies"). При этом в исследуемые пробы включалась флуоресцентная метка в составе Cy5-СТР, а контролем с Cy3-меткой выступала референсная РНК – Universal Human Reference RNA ("Agilent Technologies"). Для предотвращения адгезии малых количеств ДНК к стенкам пробирки и имитации стандартных условий синтеза к пробам добавлялась плазмидная ДНК pUC19 (вектор, не имеющий T7-промотора) в количестве 500 нг с добавлением 5% (по массе) смеси векторов pT7CFE1-CHis и pGEM-T-Easy. Первый из данных векторов включал кДНК гена *MBP* (myelin basic protein), а второй фрагмент гена *TP53* человека. Также согласно стандартному протоколу в пробы вносили соответствующие количества РНК из набора Agilent RNA Spike-In Kit.

Для гибридизации использовали микрочипы G4112F ("Agilent Technologies"), предназначенные для анализа экспрессии генов в объеме полного генома человека. Флуоресцентно меченые пробы вносились в количестве 600 нг, каждое разведение наносилось на микрочипы дважды. Сканирование проводили в двухцветном режиме на сканере G2505C при 100% чувствительности по каждому из каналов.

Первичную статистическую обработку данных производили с помощью программного обеспечения Feature Extraction 10.10.1.1 ("Agilent Technologies"). Для дальнейшего анализа и построения диаграмм использовался показатель red processed signal.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Подготовка проб для транскриптомного анализа на микрочипах фирмы "Agilent Technologies" предполагает несколько шагов: обратная транскрипция, синтез кРНК по матрице кДНК, очистку проб, дефрагментацию и гибридизацию. Второй этап представляет собой многократную амплификацию с помощью T7-РНК-полимеразы с одновременным включением флуоресцентной метки. Это становится возможным за счёт того, что на первом этапе используется праймер, включающий помимо поли-А последовательности также T7-промотор. Таким образом, можно предположить, что ферменты стандартного набора для пробоподготовки позволят обеспечить синтез любой последовательности, имеющей в своём начале область T7-промотора.

Для доказательства факта, что подобная конструкция может служить матрицей для синтеза

кРНК, необходимо также оценить зависимость между концентрацией такой ДНК с фрагментом определённого гена и интенсивностью флуоресцентного сигнала на чипе от комплементарного данному гену зонда.

В первую очередь был определен диапазон, в котором может варьировать количество молекул в таком эксперименте. По данным секвенирования мРНК одиночных клеток [8], общее количество транскриптов в клетке составляет от 50000 до 300000. В то же время известно, что большая часть генов экспрессируется в пределах от одной до десятков копий на одну клетку [9, 10]. Таким образом, для анализа желательно использовать серию разведений от единиц до сотен копий индивидуального гена. Для наиболее полной картины и достижения насыщения максимальное количество в наших экспериментах составило 1000 копий, в то время как минимальным порогом стало наличие 1 копии гена на клетку.

Для примерного расчета нужного количества ДНК-матрицы было решено исходить из нескольких условностей. Первое допущение – для проведения синтеза кДНК с последующим синтезом меченой кРНК на одну реакцию берется 500 нг общей РНК. Второе допущение – стандартная эффективность выделения РНК набором Agilent Total RNA Isolation Mini Kit принимается за 100%. Согласно протоколу выделения, данное количество (500 нг) РНК из клеточной линии 293НЕК получается примерно из 50 тысяч клеток. Таким образом, при наличии в каждой клетке 1 молекулы определенной мРНК, образец для включения флуоресцентной метки должен содержать порядка 50 тыс молекул. Дальнейшие концентрации с шагом 10, 50, 100, 500 и 1000 молекул на одну клетку рассчитывались экстраполяцией исходной цифры.

Реакционную смесь с каждым из разведений обрабатывали согласно стандартному протоколу фирмы "Agilent Technologies", начиная с этапа синтеза кДНК. В данном случае этот этап мог бы быть пропущен, поскольку матрица уже представляла собой молекулы ДНК с T7 промотором, однако он необходим для синтеза контрольных РНК набора Agilent RNA Spike-In Kit. Следует отметить также, что в предварительном эксперименте (данные не приведены) синтез флуоресцентно меченой кРНК успешно осуществлялся сразу с плазмидной ДНК.

В предварительном эксперименте была использована смесь плазмид в концентрации примерно по 30 нг каждого типа без этапа синтеза кДНК. После гибридизации была зарегистрирована флуоресценция как для плазмид с полноразмерными ORF для экспрессии цитохромов (см. раздел Методика), так и специально созданных конструкций с фрагментами генов *ROCK1*, *SMAD2* и *SMAD4*. В то же время было отмечено, что для плазмид с полноразмерными генами уровень флуоресцентного сигнала на микрочипе был намного ниже. Подобная зависимость наблюдалась и в эксперименте с полноценной пробоподготовкой, поэтому для удобства представления результатов эти группы генов показаны на отдельных графиках.

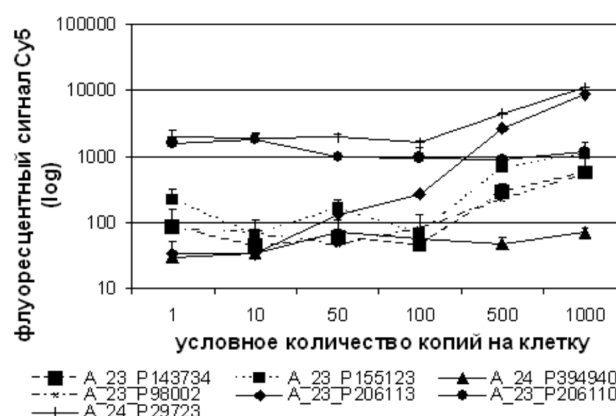
Также стоит отметить, что на микрочипе каждому из полноразмерных генов, кроме гена *POR*, соответствуют по два разных зонда (обозначены разными идентификационными номерами Agilent) (таблица). Как видно из рисунка 1, для большей части зондов заметный рост сигнала по мере увеличения дозы матрицы в пробе начинается с уровня 100 молекул на одну клетку. Исключением являются лишь два зонда A\_24\_P394940 и A\_23\_P206110, практически не имеющие положительной динамики. Здесь, однако, следует отметить, что зонд с идентификатором A\_23\_P206110 предназначен для гибридизации с зоной транскрипта гена *CYP1A2*, выпадающей из ORF данного гена. Таким образом, в данной системе для него нет подходящей матрицы и уровень флуоресценции не должен изменяться с увеличением дозы гена. По всей видимости, наблюдаемый сигнал является результатом неспецифической гибридизации, поскольку и в предварительном эксперименте с заведомо высокой концентрации гена *CYP1A2* с данного зонда наблюдался такой же уровень абсолютного сигнала ( $\approx 2000$ ).

**Таблица.** Гены, используемые в конструкциях для гибридизации и соответствующие им идентификаторы Agilent

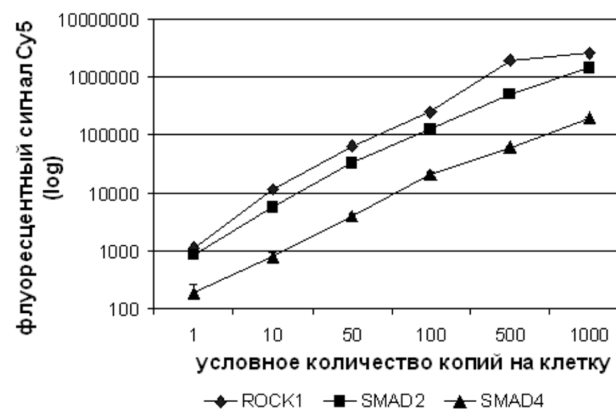
Идентификатор зонда на микрочипе Agilent	Ген	Расстояние от T7-промотора до точки гибридизации с зондом (нуклеотидов)
A_23_P143734	<i>CYP2D6</i>	972
A_23_P155123	<i>CYP2D6</i>	1262
A_24_P394940	<i>CYP2E1</i>	831
A_23_P98002	<i>CYP2E1</i>	1467
A_23_P206113	<i>CYP1A2</i>	1037
A_23_P206110	<i>CYP1A2</i>	нет гибридизации
A_24_P29723	<i>POR</i>	1940
A_23_P130429	<i>ROCK1</i>	107
A_32_P109002	<i>SMAD2</i>	113
A_23_P27346 (10 повторов)	<i>SMAD4</i>	95

Более отчетливо связь “количество ДНК – флуоресцентный сигнал” проявляется для плазмидных конструкций, несущих короткий фрагмент гена, соответствующий заранее выбранному зонду на микрочипе. На рисунке 2 представлены данные по генам *ROCK1*, *SMAD2* и *SMAD4*. Кроме того, что абсолютные величины сигнала намного превосходят аналогичные показатели генов из первой группы, здесь также становится заметна индивидуальная зависимость сигнала от концентрации. То есть при одинаковой концентрации ДНК в исходной пробе флуоресцентный сигнал от зондов разных генов может различаться на порядок, что видно на примере сравнения

показателей *ROCK1* и *SMAD4*. Можно предположить, что такая разница в интенсивности сигнала обусловлена разной степенью включения флуоресцентной метки в индивидуальные последовательности. Данный эффект был описан в работе Hesse с сотр [11], когда на примере двух генов было показано, что степень включения метки в одном случае составляет 4,6, а в другом 1,6 на молекулу кДНК. При этом измерение проводилось в высокочувствительной системе для детекции очень малых количеств меченой кДНК на микрочипах.



**Рисунок 1.** График зависимости сигнала на микрочипе от количества копий плазмидной ДНК в исходной пробе. Количество копий дано в условном пересчете на одну клетку. Приведены данные по индивидуальным зондам для генов, представленных в виде полноразмерных ORF.



**Рисунок 2.** График зависимости сигнала на микрочипе от количества копий плазмидной ДНК в исходной пробе. Количество копий дано в условном пересчете на одну клетку. Приведены данные по индивидуальным зондам для генов, представленных в плазмидных конструкциях в виде коротких фрагментов.

Что касается заметной разницы в абсолютном сигнале от конструкций с короткими участками и полноразмерными генами (сравнение рисунков 1 и 2), наиболее вероятной причиной является эффект, описанный в работе Kakuhata и соавторов [12]. На примере коммерческих чипов G4131F фирмы “Agilent Technologies” было показано, что по мере удаления местоположения зонда от 3'-конца гена (то есть участка полиаденилирования, с которого

затем начинается синтез кРНК) существенно снижается и интенсивность флуоресценции соответствующего спота на чипе. Как видно из таблицы, в нашем случае все полноразмерные гены имели значительное удаление от точки начала синтеза флуоресцентно меченой цепи кРНК. Возможно также, что при низких концентрациях матрицы данный эффект проявляется заметнее, поскольку до условного уровня 100 молекул на клетку почти не проявлялась зависимость между концентрацией матрицы и интенсивностью сигнала (рис. 1).

Из анализа графиков на рисунках 1 и 2 также следует вывод о том, что при использованных в эксперименте концентрациях матрицы точка насыщения для данных генов не была достигнута. При этом следует отметить, что синтез флуоресцентно меченой кРНК проводился в смоделированных условиях, без учета конкурентного процесса с других молекул РНК реальной биологической пробы. По всей видимости, в условиях реального эксперимента интенсивность абсолютного сигнала с зондов на эти же гены была бы несколько ниже наблюдаемой. Таким образом, выбранные условные концентрации матрицы на основе ДНК вполне могут отражать реальные соотношения молекул в биологической пробе, не превышая при этом динамический диапазон системы детекции.

## ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

ДНК-конструкция с участком гена под контролем Т7-промотора может быть использована как полноценная матрица для синтеза флуоресцентно меченой РНК, пригодной для гибридизации на микрочипах. При этом намного эффективнее синтез проходит с коротких участков гена, расположенных непосредственно после Т7-промотора. Преимущество такой матрицы заключается в простоте изготовления, очистки, большей стабильности по сравнению с РНК и достаточно простому контролю количества молекул по концентрации. Также подобная матрица универсальна, поскольку может включать участок с любой синтезированной последовательностью.

По всей видимости, область применения такой конструкции на основе плазмиды сравнима с использованием в гибридизационной транскриптомике РНК, синтезированной *in vitro*. В случае разработки оригинальных микрочипов

можно заранее вводить на чип определенное количество зондов для такой референсной системы. Плазмиды с участками, комплементарными этим зондам, могут наноситься в виде смеси в разных концентрациях, что позволит осуществлять как мониторинг качества пробоподготовки и гибридизации, так и проводить нормализацию и стандартизацию данных.

*Работа частично поддержана грантом РНФ №14-25-00132.*

## ЛИТЕРАТУРА

1. Dudley A.M., Aach J., Steffen M.A., Church G.M. (2002) Proc. Natl. Acad. Sci. USA, **99**, 7554-7559.
2. Carter M.G., Sharov A.A., VanBuren V., Dudekula D.B., Carmack C.E., Nelson C. et al. (2005) Genome Biology, **6**, R61.
3. Kanno J., Aisaki K., Igarashi K., Nakatsu N., Ono A., Kodama Y., Nagao T. (2006) BMC Genomics, **7**, p. 64.
4. Hill A.A., Brown E.L., Whitley M.Z., Tucker-Kellogg G., Hunter C.P., Slonim D.K. (2001) Genome Biology, **2**, RESEARCH0055.
5. van de Peppel J., Kemmeren P., van Bakel H., Radonjic M., van Leenen D., Holstege F.C. (2003) EMBO Rep, **4**, 387-393.
6. Kakuha R., Watanabe M., Yamamoto T., Akamine R., Yamazaki N., Kataoka M. et al. (2007) J. Biochem. Biophys. Methods, **70**, 755-760.
7. Yang I.V. (2006) Methods Enzymol., **411**, 50-63.
8. Marinov G.K., Williams B.A., McCue K., Schroth G.P., Gertz J., Myers R.M., Wold B.J. (2014) Genome Res, **24**(3), 496-510.
9. Wodicka L., Dong H., Mittmann M., Ho M.H., Lockhart D.J. (1997) Nat. Biotechnol., **15**, 1359-1367.
10. Zenklusen D., Larson D.R., Singer R.H. (2008) Nat. Struct. Mol. Biol., **15**, 1263-1271.
11. Hesse J., Jacak J., Kasper M., Regl G., Eichberger T., Winklmayr M., Aberger F., Sonnleitner M., Schlapak R., Howorka S., Muresan L., Frischauf A.M., Schutz G.J. (2006) Genome Res., **16**, 1041-1045.
12. Kakuha R., Watanabe M., Yamamoto T., Obana E., Yamazaki N., Kataoka M., Ooe T., Baba Y., Hori T., Shinohara Y. (2008) J. Biochem. Biophys. Methods, **70**, 926-931.

Поступила: 07. 09. 2016.  
Принята к печати: 25. 10. 2016.

**A POSSIBILITY TO USE THE DNA-BASED PROBES AS INTERNAL STANDARDS  
FOR AGILENT TECHNOLOGIES MICROARRAY TRANSCRIPTOMIC ANALYSIS**

*L.K. Kurbatov, V.G. Zgoda*

Institute of Biomedical Chemistry,  
10 Pogodinskaya str., Moscow, 119121 Russia; e-mail: kurbatovl@mail.ru

Accuracy of the microarray technology results is raised by using the multi-stage normalization of results. One of the principal requirements of such normalization is usage of internal standards. The routine Agilent microarray-based gene expression analysis protocol utilizes a Spike-In Kit during preparation of the samples representing a mixture of RNA fragments in different ratios. RNA probes which were synthesized in vitro conditions could be also used to establish how the magnitude of the fluorescent signal reflects the presence of RNA in the sample. A significant disadvantage of this type of standards is a difficulty of their production and the low RNA stability. In accordance with the Agilent protocol, the presence of the T7 promoter is necessary for the synthesis of labeled cRNA during sample preparation procedure. We hypothesized that we can successfully synthesize any RNA sequence having such type of promoter in its start position. Moreover, DNA sequence would serve as a matrix in this case. Using a set of different genes attached downstream of the T7-promoter in the plasmid DNA we have demonstrated in this study that such system can serve as a reliable template for the fluorescent labeled RNA sequence synthesis. In comparison with the routinely used internal RNA based controls, this template is stable, easy to manufacture and can be easily obtained in large quantities.

**Key words:** microarrays, plasmid DNA, T7-promoter, fluorescence level, labeled cRNA