

Стандартный образец нуклеотидной последовательности молекулы ДНК

С.С. Голубев^{*}, Ю.А. Кудеяров^{*}, А.В. Марданов^{}, П.Ю. Николаева^{*}, Н.В. Равин^{**}**

^{*} Всероссийский научно-исследовательский институт метрологической службы

^{**} Центр «Биоинженерия» РАН

Задача определения последовательности нуклеотидов в молекуле ДНК в последние годы приобрела чрезвычайную важность и актуальность, что обусловлено в первую очередь необходимостью получения информации о генетическом коде живых организмов для решения ряда важных задач геномики, генной инженерии, молекулярной диагностики, медицины, ветеринарии, сельского хозяйства и т.п. Решение этой задачи стало возможным благодаря прогрессу биологического приборостроения – автоматизации рутинных процедур, миниатюризации, объединению различных модулей в интегрированные многофункциональные системы, широкому использованию возможностей вычислительной техники, приведших к значительному увеличению производительности отдельных биологических экспериментов и выведению исследований на новый качественный уровень.

Методы расшифровки нуклеотидной последовательности ДНК, состоящей из двух комплементарных цепочек, каждая из которых состоит из большого числа нуклеотидов четырех типов: аденина (А), гуанина (G), цитозина (С) и тимина (Т), принято называть методами секвенирования (последовательного прочтения). Для решения этой задачи используются специальные приборы – секвенаторы. Почти все современные секвенаторы используют принцип «секвенирования путем синтеза», т.е. определения нуклеотидной последовательности исследуемой ДНК посредством включения отдельных нуклеотидов во вновь синтезируемую по принципу ком-

плементарности цепочку. Такой процесс аналогичен ферментативной репликации ДНК в клетке, осуществляемой ДНК-полимеразами.

В настоящее время одним из распространенных методов расшифровки последовательностей ДНК является пирофосфатное секвенирование или пиросеквенирование. Принцип действия пиросеквенаторов, а также основные понятия и определения, используемые в этой области биохимии, детально рассмотрены в работе [1], что освобождает от необходимости еще раз останавливаться на этих вопросах.

Поскольку указанная деятельность осуществляется в сфере государственного регулирования в области обеспечения единства измерений, то особую важность приобретает разработка и осуществление метрологического обеспечения соответствующих измерений, в частности, секвенирования. Проблема становится особенно актуальной в связи с большим затратным характером процедуры секвенирования и высокими требованиями к достоверности результатов определения нуклеотидной последовательности генома, обусловленными не в последнюю очередь его возможными социальными последствиями (например, последовательности индивидуального генома человека могут указывать на предрасположенность к различным заболеваниям). Важным элементом такого метрологического обеспечения являются испытания пиросеквенаторов для целей утверждения типа и внесения их в Государственный реестр средств измерений.

Ситуация, однако, осложняется тем, что секвенаторы не являются средствами измерений в традиционном смысле, поскольку не воспроизводят и не хранят измеряемую величину. Это скорее устройства с измерительными функциями. Кроме того, в биологии практически отсутствует какая-либо эталонная база и такие атрибуты метрологического обеспечения как методики измерений и методики поверки (калибровки). Выход из создавшегося положения был предложен в работах [1-2].

Так в работе [2] предлагается рассматривать генетическую последовательность молекулы ДНК как новую величину для метрологии и ставится задача нахождения для нее необходимой ниши. Отмечается, что, в первую очередь, необходима разработка стандартного образца (СО) «эталонной» последовательности нуклеотидов, позволяющего проводить поверку и калибровку секвенаторов. Этот стандартный образец должен характеризоваться тем, что в нем должна быть достоверно известна последовательность нуклеотидов ДНК. В качестве стандартного образца предлагается использовать такой «эталон» – один из фрагментов последовательности, обозначаемой в биологии плазмидой pUC18 (Yanisch-Perron, C., et al., Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors, *Gene*, 33, 103-119, 1985). Этот фрагмент состоит из 271 пары нуклеотидов, он очень хорошо изучен и многократно прочитан в различных лабораториях мира. Отличительная его особенность заключается в свойстве природной «фундаментальности», поскольку благодаря исключительно высокой точности копирования генетической информации в природе последовательность нуклеотидов одинакова у всех экземпляров плазмиды pUC18. Кроме того, эта цепочка содержит участки из одинаковых нуклеотидов вплоть до пяти штук подряд, что позволяет исследовать секвенаторы на правильность прочтения таких «проблемных» участков. Дело в том, что при наличии в цепи нескольких одинаковых нуклеотидов подряд из-за возможного нарушения линейной зависимости (пропорциональности) регистрируемой ПЗС – матрицей (матрицей из приборов с зарядовой связью) интенсивности светового сигнала от числа включенных в растущую цепь нуклеотидов возможно неправильное прочтение секвенатором таких участков особенно тогда, когда друг за другом следуют три-четыре одинаковых нуклеотида и более.

В качестве метрологических характеристик секвенаторов в работе [1] предлагается рассматривать следующие характеристики прочтения:

долю ошибочно прочтенных нуклеотидов последовательности ДНК,

долю правильных прочтений всей последовательности стандартного образца, а также прочтений, содержащих различное число ошибок (1, 2, 3 и более),

долю правильных прочтений участков стандартного образца с i одинаковыми нуклеотидами подряд.

С помощью разработанного СО были успешно проведены испытания секвенатора геномного GS FLX, изготовленного фирмой Roche Diagnostics GmbH (Германия) [1].

Вместе с тем, развитие техники пиросеквенирования ставит вопрос о совершенствовании стандартных образцов в направлении расширения их функциональных возможностей и повышения надежности определения метрологических характеристик секвенаторов. Речь идет, прежде всего, об удлинении цепочки достоверно прочитанных нуклеотидов в СО. Желательно также, чтобы в такой цепочке были участки, состоящие из одинаковых нуклеотидов (числом более пяти, например из шести – семи), что должно повысить достоверность проверки возможностей секвенаторов на правильное прочтение таких «проблемных» участков. Кроме того, наличие СО с указанными свойствами позволит поставить в повестку дня вопрос о разработке референтной методики секвенирования молекул ДНК.

Таким образом, в связи с указанными причинами, а также в связи с модернизацией анализатора геномного GS FLX, увеличившей среднюю длину прочтения индивидуальной реакции с 350-400 до 500-800 нуклеотидов, и необходимостью усовершенствования механизма определения оши-

бок при чтении последовательности ДНК посредством пиросеквенаторов возникла необходимость разработки нового стандартного образца, предназначенного для проведения работ по поверке и калибровке пиросеквенаторов.

Федеральным государственным унитарным предприятием «Всероссийский научно-исследовательский институт метрологической службы» (ФГУП «ВНИИМС») совместно с Центром «Биоинженерия» РАН в рамках НИОКР «Разработка эталонной последовательности нуклеотидов в целях создания эталонной референтной методики измерений геномной последовательности» в 2012 г. была проведена разработка стандартного образца фрагмента плазмиды pUC18, состоящего из 717 пар нуклеотидов.

Предлагаемая статья посвящена краткому изложению результатов этой работы.

Для разработки стандартного образца были сформулированы следующие требования:

стандартный образец должен быть получен из стабильной, не склонной к мутациям, легко воспроизводимой молекулы ДНК;

фрагмент молекулы ДНК, выбранный в качестве стандартного образца должен состоять из 500–800 нуклеотидов;

последовательность стандартного образца должна содержать участки, состоящие из 2-7 одинаковых нуклеотидов подряд;

массовая концентрация ДНК в водном растворе в диапазоне должна находиться в пределах от 32 до 83 нг/мкл.

Границы относительной погрешности измерения массовой концентрации ДНК в растворе не должны превышать $\pm 15\%$ при доверительной вероятности $P = 0,95$.

На основе сформулированных требований было разработано техническое задание на разработку стандартного образца эталонной последовательности нуклеотидов.

Руководствуясь требованиями, изложенными в техническом задании, был сделан вывод о целесообразности получения стандартного образца из плазмиды pUC18. Напомним, что плазмиды – внехромосомные автономно реплицирующиеся генетические элементы клетки, представляющие собой преимущественно кольцевые замкнутые молекулы ДНК.

В разработанном стандартном образце должны быть представлены участки, содержащие повторяющиеся нуклеотиды в следующем соотношении:

участки последовательности стандартного образца, содержащие 2 одинаковых нуклеотида подряд – 103 участка;

участки последовательности стандартного образца, содержащие 3 одинаковых нуклеотида подряд – 21 участок;

участки последовательности стандартного образца, содержащие 4 одинаковых нуклеотида подряд – 8 участков;

участки последовательности стандартного образца, содержащие 5 одинаковых нуклеотидов подряд – 4 участка;

участки последовательности стандартного образца, содержащие 6 одинаковых нуклеотидов подряд – 1 участок;

участки последовательности стандартного образца, содержащие 7 одинаковых нуклеотидов подряд – 2 участка.

Центром «Биоинженерия» РАН была разработана методика приготовления стандартного образца эталонной последовательности. В этой методике изложена процедура получения плазмиды pUC18, являющейся материалом разрабатываемого стандартного образца, и процедура получения стандартного образца фрагмента плазмиды pUC18, состоящего из 717 пар нуклеотидов из плазмиды pUC18. В свою очередь, плазмиду pUC18 получают из специального штамма бактерий кишечной палочки *Escherichia coli*. Размеры статьи не позволяют в деталях изложить методику изготовления СО. Эта методика изложена в работах [3-4]. Здесь мы остановимся только на рассмотрении лишь некоторых особенностей этой методики.

Процедура приготовления стандартного образца фрагмента плазмиды pUC18, состоящего из 717 пар нуклеотидов, включает в себя следующие этапы:

обработка препарата ДНК плазмиды pUC18 рестриктазами PvuI и Eco571

очистка препарата и получение фрагмента, состоящего из 717 пар нуклеотидов

Метод получения СО основан на расщеплении последовательности pUC18 специальными ферментами рестрикции - рестриктазами. Рестриктаза это бактериальный фермент, распознающий специфическую короткую (длиной 4 - 6 нуклеотидов) последовательность в двухцепочечной ДНК и расщепляющий, «разрезающий» непрерывную цепочку ДНК только на этом участке (сайте). Для получения СО использовались рестриктазы PvuI и Eco571. Эти рестриктазы имеют четыре сайта узнавания последовательности pUC18 и расщепляют ее на 4 фрагмента по 1075, 717, 297 и 599 пар

нуклеотидов по сайтам, расположенным в позициях 98, 1173, 1888 и 2185 соответственно (см. рис. 1).

Стандартный образец представляет собой криопробирку с водным раствором фрагмента двухцепочечной ДНК плазмиды pUC18 размером 717 нуклеотидов (каждая цепочка состоит из 717 пар нуклеотидов). Объем водного раствора стандартного образца составляет 200 мкл.

Сотрудниками ФГУП «ВНИИМС» были определены основные физико-химические свойства созданного стандартного образца эталонной последовательности, которые было необходимо установить для утверждения типа стандартного образца.

Следует отметить, что препараты ДНК могут иметь вид как высушенного осадка (твердая фаза), так и раствора (жидкая фаза). В зависимости от того, в какой фазе находится препарат, меняются условия его хранения и использования. В качестве параметра, обуславливающего условия хранения и эксплуатации стандартного образца, выступает температура. Помимо определения фазы и температуры необходимо определить плотность и объем стандартного образца, а так же такие его свойства, как цвет и прозрачность.

Для определения свойств стандартного образца были выбраны методы оценки, представленные в таблице 1.

Таблица 1

Исследуемое свойство	Метод исследования
1	2
Фаза (жидкая/твердая/)	метод экспертной оценки
Объем	измерения с использованием дозаторов с переменным объемом дозирования по

	ТУ 9452-002-33189998-2002
Плотность	расчетный метод (масса/объем)
Цвет	метод экспертной оценки
Прозрачность	метод экспертной оценки
Температура хранения	метод экспертной оценки

Исполнителями уже упомянутой НИОКР были проведены экспериментальные исследования указанных выше свойств разработанного стандартного образца.

В ходе исследований были получены следующие результаты:

методом экспертной оценки было определено, что стандартный образец фрагмента плазмиды рUC18, состоящий из 717 пар нуклеотидов, представляет собой жидкую фазу (водный раствор ДНК);

объем водного раствора ДНК стандартного образца определялся посредством дозаторов с переменным объемом и составил 200 мкл;

плотность стандартного образца была определена расчетным методом, путем нахождения отношения массы стандартного образца (9,8 мкг) к его объему (200 мкл), и составила 49 нг/мкл.

Затем методом экспертной оценки было определено, что стандартный образец бесцветен и прозрачен.

По результатам экспертной проверки был также сделан вывод о том, что оптимальной для хранения стандартного образца фрагмента плазмиды рUC18, состоящего из 717 пар нуклеотидов, является температура, равная минус 20 °С и ниже.

Чтобы убедиться в том, что структура СО соответствует требованиям технического задания, и в том, что она правильно читается секвенатором, разработанный СО был подвергнут пиросеквенированию. Процедура секвенирования детально изложена в работе [1]. Отметим только, что секвенирование фрагментов ДНК происходит в каждой лунке пикотитровальной ячейки, что, в свою очередь, фиксируется с помощью оптической системы, выполненной на основе ПЗС – матрицы. Полученные изображения передаются на управляющий компьютер, обрабатывающий полученные данные.

В геномном анализаторе GS FLX определение нуклеотидных последовательностей осуществляется одновременно в миллионах микроячеек, находящихся на пикотитровальной пластине. Таким образом, один и тот же фрагмент ДНК, может быть многократно прочитан, что позволяет построить для него единую обобщённую (так называемую консенсусную) последовательность нуклеотидов. Отдельные прочтения одного и того же участка ДНК выравниваются относительно друг друга, исходя из интенсивности сигналов в момент протекания через камеру того или иного нуклеотида. Затем соответствующие сигналы усредняют и записывают полученную последовательность.

Определение последовательности нуклеотидов в полученной молекуле ДНК проводится в автоматическом режиме в несколько этапов.

Первый этап обработки результатов испытаний проводится на компьютерном кластере Titanium (Roche). При этом финальная обработка сигнала, запускается автоматически сразу после окончания прочтений фрагментов пиросеквенатором. Данные сохраняются в файле формата .SFF и представляют собой набор «прочтенных» последовательностей нуклеотидов СО, т.е. последовательность букв, которыми обозначаются нуклеотиды.

Второй этап обработки результатов измерений обусловлен тем, что последовательность нуклеотидов стандартного образца анализируется в одних случаях как прямая, а в других – как комплементарная. При этом, естественно, возникает необходимость их сопоставления. Этот этап обработки результатов исследований проводится следующим образом.

Прежде всего, находятся участки прочитанных последовательностей с минимальным перекрытием (overlap) длиной в 100 нуклеотидов и идентичностью > 99%. Считается, что чтения перекрываются, если существует выравнивание от конца одного чтения до конца другого, при этом случаи, когда одно чтение полностью включено в другое, не рассматриваются (рис. 2).

Выравнивания строятся по алгоритму Смита – Ватермана, который сводится к следующей последовательности действий.

Для двух последовательностей $A = a_1, a_2, \dots, a_n$ и $B = b_1, b_2, \dots, b_m$ заполняется матрица переходов H размерностью $(n + 1) \times (m + 1)$, при этом значения элементов верхней строки и левого столбца матрицы переходов приравниваются нулю, т.е. $H_{0,0} = H_{i,0} = H_{0,j} = 0$.

Все остальные элементы считаются по формуле:

$$H_{i,j} = \max \begin{cases} H_{i-1,j-1} + s(a_i, b_j) \\ H_{i-1,j} - g \\ H_{i,j-1} - g \\ 0 \end{cases},$$

где g - штраф за вставку ($g = 3$);

$$s(a_i, b_j) = 2, \quad \text{если } a_i = b_j,$$

$$s(a_i, b_j) = -3, \quad \text{если } a_i \neq b_j.$$

Для каждого положительного элемента $H_{i,j}$ запоминается направление перехода. Переходы осуществляются по следующему правилу.

Если $H_{i,j} = H_{i-1,j} + g$, то переход в элемент $H_{i,j}$ осуществляется из элемента $H_{i-1,j}$;

если $H_{i,j} = H_{i,j-1} + g$, то переход осуществляется из элемента $H_{i,j-1}$;

если $H_{i,j} = H_{i-1,j-1} + s(a_i, b_j)$ — из элемента $H_{i-1,j-1}$.

Чтобы получить оптимальное локальное выравнивание, из ячейки с максимальным значением строится путь по направлениям перехода до ячейки с нулевым значением.

Если значение ячейки в пути получено из диагонального перехода, то в выравнивании приравниваются соответствующие пары нуклеотидов, т.е. нуклеотид и вставка. Указанные правила иллюстрируются рис. 3.

Далее берется одна пара перекрывающихся чтений и строится предварительный контиг (1-вое чтение + невыравненный остаток 2го) (см. рис. 4). Напомним, что контиг это непрерывная последовательность нуклеотидов, образуемая группой из нескольких перекрывающихся секвенированных последовательностей ДНК. Относительно этого предварительного контига выравниваются все оставшиеся чтения по отдельности, при этом для каждой позиции предварительного контига подчитывается число совпадений и число различных несовпадений, и в итоговом консенсусе выбирается наиболее часто встречаемый вариант.

Третий и заключительный этап обработки прочтения последовательности нуклеотидов СО заключается в определении окончательного результата прочтения.

Результаты, полученные после окончания второго этапа обработки сопоставляются, при этом сравниваются буквы в каждой позиции. Для записи результата прочтения для каждой из позиций последовательности СО выбирается буква, встретившаяся наибольшее количество раз. (Например: при 10 измерениях 3-ая позиция в каждой прочитанной последовательности имеет следующий вид: AAACAAGAAA, следовательно в результате измерений аттестованного значения последовательности нуклеотидов стандартного образца 3-ая позиция последовательности записывается буквой А). Результатом прочтения последовательности нуклеотидов СО является так называемая консесусная последовательность, которая изображена на рис. 5.

Важным этапом работы было утверждение типа разработанного СО. С этой целью была разработана «Методика испытаний», целью которых является решение вопроса о соответствии технических и метрологических характеристик СО требованиям технического задания.

Испытаниями предусмотрено проведение следующих процедур:

определение метрологических характеристик СО;

исследование его стабильности;

исследование однородности СО;

установление срока годности.

В соответствии с требованиями «Методики испытаний» значение массовой концентрации ДНК в водном растворе должно находиться в пределах от 32 до 83 нг/мкл, а границы относительной погрешности измерений концентрации при доверительной вероятности $P=0,95$ не должны превышать $\pm 15\%$.

Метод определения метрологических характеристик выбирался в соответствии с ГОСТ 8.315-97 [5], п. 5.4, разд. а), т.е. с использованием СО утвержденного типа. При этом определение массовой концентрации ДНК в водном растворе осуществлялось с применением модифицированной системы капиллярного электрофореза G1600А Фирма «Agilent Technologies», США. Количественные измерения проводились методом градуировки, при этом для градуировки использовался государственный стандартный образец (ГСО) 9931-2011.

Таким образом, системой капиллярного электрофореза одновременно проводились измерения массовой концентрации испытываемого стандартного образца фрагмента плазмиды pUC18, состоящего из 717 пар нуклеотидов, и ГСО фрагмента плазмиды pUC18 длиной 271 нуклеотид утвержденного типа, используемого для градуировки секвенатора. Стоит отметить, что массовая концентрация ДНК в растворе определялась как количество вещества ДНК в единице объема, поэтому от количества нуклеотидов в последовательностях отдельных молекул ДНК она не зависит.

Для проведения измерений изготавливались специальные чипы с анализируемым стандартным образцом и ГСО. Процедура приготовления чипов подробно изложена в «Методике испытаний».

Проводилось 10 измерений массовой концентрации ДНК в водном растворе как анализируемого СО, так и ГСО.

Обработка результатов экспериментального исследования подразумевает под собой расчет значения массовой концентрации ДНК в водном растворе, а так же определение относительной погрешности измерений.

В «Методике испытаний» использован метод, рекомендуемый методикой МИ 3174-2009 [6]. Согласно этому методу для определения аттестованного значения анализируемого СО необходимо рассчитать отноше-

ния результатов измерений анализируемого СО и ГСО, т.е. провести расчет по формуле

$$g_i = \frac{I_{ai}}{I_{oi}},$$

где I_{ai} – результаты измерений массовой концентрации ДНК в водном растворе анализируемого СО, I_{oi} – результаты измерений массовой концентрации ДНК в водном растворе ГСО

Затем проводился расчет среднего значения отношений результатов измерений анализируемого СО и ГСО

$$\bar{g} = \frac{1}{n} \sum_{i=1}^n g_i$$

где n – число измерений, равное 10.

Стандартное (среднее квадратическое) отклонение среднего отношения результатов измерений анализируемого СО и ГСО рассчитывалось по формуле

$$S_{\bar{g}} = \sqrt{\frac{1}{n \cdot (n-1)} \sum_{i=1}^n (g_i - \bar{g})^2}$$

Аттестованное значение массовой концентрации ДНК в водном растворе СО определялось по формуле:

$$\eta = \bar{g} \cdot \eta_o,$$

где η_o – аттестованное значение концентрации ДНК в водном растворе ГСО, равное 49,1 нг/мкл..

Границы абсолютной погрешности измерений массовой концентрации ДНК в водном растворе СО рассчитывались по формуле

$$\Delta\eta = 2 \cdot \sqrt{\eta_0^2 \cdot S_g^2 + \frac{\Delta(\eta_0)^2}{3}},$$

где $\Delta(\eta_0)$ – абсолютная погрешность концентрации ДНК в водном растворе ГСО ($\Delta(\eta_0) = \frac{\delta\eta_0 \cdot \eta_0}{100\%}$), $\delta\eta_0$ – относительная погрешность измерений концентрации ДНК в водном растворе ГСО, равная 11%.

Относительная погрешность измерения массовой концентрации ДНК в водном растворе СО рассчитывалась по формуле

$$\delta\eta = \frac{\Delta\eta}{\eta} \cdot 100\%$$

Уже отмечалось, что измерения были выполнены на системе капиллярного электрофореза G1600A (номер в Госреестре СИ № 15720-96), при этом результаты измерений прослеживаются к ГСО № 9931-2011. Результаты измерений и расчетов, проведенных в соответствии с «Методикой испытаний», массовой концентрации ДНК в растворе СО с занесены в соответствующие протоколы испытаний (см. таблицу 2).

Таблица 2

№ измерения	Результаты измерений массовой концентрации анализируемого СО (I_{ai}), нг/мкл	Результаты измерений массовой концентрации ГСО (I_{0i}), нг/мкл	Отношения результатов измерений анализируемого СО и ГСО, (v)
1	49,7	55,2	0,9
2	50,6	52,8	1,04
3	48,9	50,6	1,07
4	49,2	53,4	1,08
5	50,1	53,8	1,07

6	49,8	54,1	1,08
7	48,7	52,5	1,08
8	49,5	51,4	1,06
9	50,2	50,9	1,01
10	49,6	54,2	1,09
Среднее значение отношений результатов измерений анализируемого СО и ГСО (\bar{g})		1,05	
Стандартное отклонение среднего отношения результатов измерений анализируемого СО и ГСО ($S_{\bar{g}}$)		0,02	
Аттестованное значение массовой концентрации ДНК в водном растворе СО (η), нг/мкл		52,2	
Относительная погрешность измерений массовой концентрации ДНК в водном растворе СО ($\delta\eta$), %		7	

Следует отметить, что определение аттестованных значений метрологических характеристик испытываемого СО проводилось в соответствии с аттестованной методикой (методом) измерений [3], разработанной в 2011 г. при создании комплекса метрологического обеспечения процессов пирофосфатного секвенирования ДНК. В соответствии с этой методикой показателем точности прочтения последовательности ДНК методом пирофосфатного секвенирования является доля ошибочно прочтенных нуклеотидов, составляющая 2% от количества нуклеотидов в анализируемой последовательности при одном прочтении. Поскольку при испытании стандартного образца проводится порядка полутора тысяч одновременных прочтений, то доля неверно прочтенных нуклеотидов в окончательном результате испытаний оказывается ничтожно малой, из чего можно сделать

вывод о том, что результат проведенных испытаний оказывается практически стопроцентно достоверным, а это означает, что этим показателем точности при записи результатов можно пренебречь.

Для контроля стабильности СО проводились измерения массовой концентрации ДНК в водном растворе в соответствии с «Методикой испытаний», а также контроль последовательности нуклеотидов и количества пар нуклеотидов в последовательности СО.

Стоит отметить, что для определения стабильности экземпляра СО необходимо и достаточно установить влияние условий хранения на стабильность его характеристик.

Измерения массовой концентрации ДНК в водном растворе проводились в течение 1 года с интервалом в 30 дней при условиях хранения минус $(20 \pm 0,5)$ °С.

Контроль стабильности последовательности нуклеотидов и количества пар нуклеотидов в последовательности СО проводился в течение 2 лет с интервалом 6 месяцев при условиях хранения минус $(20 \pm 0,5)$ °С.

Расчет срока годности проводится в соответствии с п. 5 Р 50.2.058-2007 [7].

Статистически значимого изменения за период исследования стабильности не обнаружено.

В случае, если изложенные выше требования соблюдены, а последовательность нуклеотидов и количество пар нуклеотидов в последовательности СО остается неизменным, то срок годности СО принимают равным двум годам.

По материалам проведенных испытаний для целей утверждения типа разработанный стандартный образец фрагмента плазмиды pUC18, состоящий из 717 пар нуклеотидов, был занесен в Государственный реестр стандартных образцов.

Авторы благодарят О.Н. Мелкову за помощь в подборе материала и написании статьи.

Литература

1. Голубев С.С., Кононогов С.А., Кудеяров Ю.А., Марданов А.В., Николаева П.Ю., Равин Н.В., Скрябин К.Г. Метрологическое обеспечение секвенирования молекул ДНК // Измерительная техника, № 3, 2012, с.с. 64-72.
2. Голубев С. С., Кононогов С.А., Равин Н.В, Скрябин К.Г. Выбор эталона для определения нуклеотидных последовательностей молекул ДНК // Измерительная техника. 2011. № 12. с. 45–47.
3. Методика (метод) измерений. Определение нуклеотидных последовательностей ДНК с использованием анализатора геномного CS FLX ФР.1.39.2011.10815. Свидетельство об аттестации № 01.00225-2008.009/001-11 от 03.10.2011.
4. Отчет о НИОКР
5. ГОСТ 8.215-97. ГСИ. Стандартные образцы состава и свойств веществ и материалов. Основные положения.
6. МИ 3174-2009 ГСИ. Установление прослеживаемости аттестованных значений стандартных образцов.
7. Р 50.2.058-2007

Рисунки

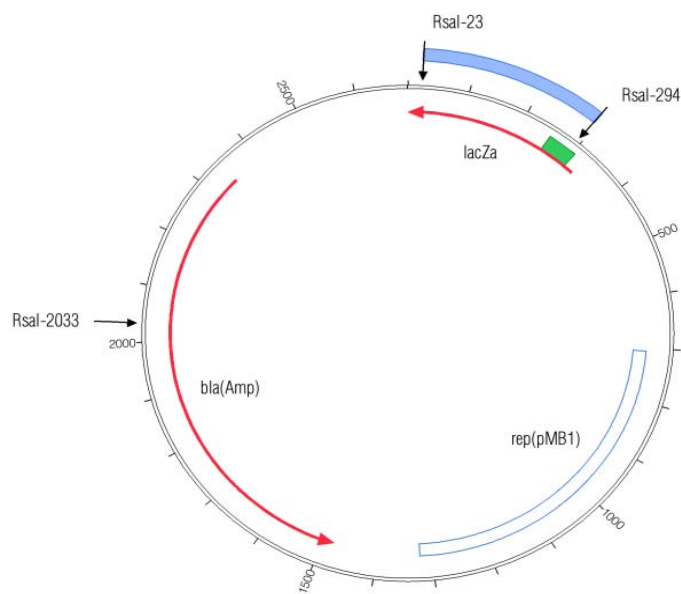


Рис. 1. Рестрикционная карта плазмиды pUC18.



Рис.2. Перекрытия участков последовательности нуклеотидов

		A	T	C	T	G	A
	0	0	0	0	0	0	0
C	0	0	0	2	0	0	0
T	0	0	2	0	4	1	0
A	0	2	0	0	1	1	3
G	0	0	0	0	0	3	0
A	0	2	0	0	0	0	5
A	0	2	0	0	0	0	2

Рисунок 3. Пример матрицы переходов для последовательностей
 CT-GA
 CTAGA



Рис. 4. Построение предварительного контига

1 AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG
GTGGTTTTTT
TTTTTCTCAA CCATCGAGAA CTAGGCCGTT TGTTTGGTGG CGACCATCGC
CACCAAAAAA
61 TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC
CTTTGATCTT
ACAAACGTTC GTCGTCTAAT GCGCGTCTTT TTTTCTAGA GTTCTTCTAG
GAAACTAGAA
121 TTCTACGGGG TCTGACGCTC AGTGGAACGA AAATCACGT TAAGGGATT
TGGTCATGAG
AAGATGCCCC AGACTGCGAG TCACCTTGCT TTTGAGTGCA ATTCCCTAAA
ACCAGTACTC
181 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTAAATTA AATGAAGTT
TTAAATCAAT
TAATAGTTTT TCCTAGAAGT GGATCTAGGA AAATTTAATT TTTACTTCAA
AATTTAGTTA
241 CTAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA
GTGAGGCACC
GATTCATAT AACTCATT GAACCAGACT GTCAATGGTT ACGAATTAGT
CACTCCGTGG

301 TATCTCAGCG ATCTGTCTAT TTCGTTTCATC CATAGTTGCC TGA^{CTCCCCG}
TCGTGTAGAT

ATAGAGTCGC TAGACAGATA AAGCAAGTAG GTATCAACGG ACTGAGGGGC
AGCACATCTA

361 AACTACGATA CGGGAGGGCT TACCATCTGG CCCAGTGCT GCAATGATAC
CGCGAGACCC

TTGATGCTAT GCCCTCCCGA ATGGTAGACC GGGGTCACGA CGT^{TACTATG}
GCGCTCTGGG

421 ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG
CCGAGCGCAG

TGCGAGTGGC CGAGGTCTAA ATAGTCGTTA TTTGGTCGGT CGGCCTTCCC
GGCTCGCGTC

481 AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC
GGGAAGCTAG

TTCACCAGGA CGTTGAAATA GCGGAGGTA GGTCAGATAA TTAACAACGG
CCCTTCGATC

541 AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA
CAGGCATCGT

TCATTCATCA AGCGGTCAAT TATCAAACGC GTTGCAACAA CGGTAACGAT
GTCCGTAGCA

601 GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTC^{CCAAC}
GATCAAGGCG

CCACAGTGCG AGCAGCAAAC CATAACGAAG TAAGTCGAGG CCAAGGGTTG
CTAGTTCCGC

661 AGTTACATGA TCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC
CTCCGAT

TCAATGTACT AGGGGGTACA ACACGTTTTT TCGCCAATCG AGGAAGCCAG
GAGGCTA

Рис.5. Последовательность нуклеотидов в стандартном образце.